

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 16:05:16 ; Search time 6071.42 Seconds
(without alignments)
11868.013 Million cell updates/sec

Title:	US-09-980-464-4_COPY_123_2015
Perfect score:	1893
Sequence:	1 atggagtcggtgcctact.....gatctgtcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 Begs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
1:  gb_est1: *
2:  gb_est2: *
3:  gb_hrc: *
4:  gb_est3: *
5:  gb_est4: *
6:  gb_est5: *
7:  gb_est6: *
8:  gb_gsa1: *
9:  gb_gsa2: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1893	100.0	2899	3	AK004737	AK004737 Mus muscu
2	1893	100.0	3098	3	AK033672	AK033672 Mus muscu
3	1857.4	98.1	2869	3	AK033082	AK033082 Mus muscu
4	1773	93.7	1896	9	AY399041	AY399041 Mus muscu
5	1355.4	71.6	3443	3	BSM001859	ALI16891 Homo sapi
6	1261	66.6	1887	9	AY398035	AY398035 Homo sapi
7	948.4	50.1	1887	9	AY399040	AY399040 Pan trogl
8	925.4	48.9	7004	3	CR749209	CR749209 Homo sapi
9	683.6	36.1	726	6	CB248251	CB248251 UT-M-FD0
10	672	35.5	672	7	CO424332	CO424332 UT-M-HU0
11	652.4	34.5	655	7	CK781308	CK781308 UT-M-G10
12	651.6	34.4	805	4	BGI174288	BGI174288 G023345889
13	646	34.1	946	4	BG694881	BG694881 NISC_1v09
14	645.8	34.1	946	2	BF789245	BF789245 G02105144
15	640	33.8	795	6	CA319312	CA319312 UT-M-FW0
16	627.6	33.2	782	7	CN261003	CN261003 170004245
17	614.8	32.5	1059	5	BM922376	BM922376 AGENCOURT
18	605.8	32.0	1022	5	B0628868	B0628868 AGENCOURT
19	603	31.9	893	3	BF785290	BF785290 G02108490
20	596	31.0	705	7	CF726196	CF726196 UT-M-GZ0
21	586.4	31.0	1005	5	BQ067629	BQ067629 AGENCOURT
22	586.4	31.0	1029	5	B0055508	B0055508 AGENCOURT
23	578	30.5	722	4	BI143493	BI143493 G02907673
24	574.4	30.3	1012	5	BQ064231	BQ064231 AGENCOURT

25	562.8	29.7	713	2	B8625283	B8625283
26	558.8	29.5	667	7	B1648260	B1648260
27	554.8	29.3	558	8	BH116351	RPCT -24-2
28	552.2	29.2	1042	5	BM909401	AGENCYCOURT
29	547.2	28.9	939	2	BE917458	BE917458
30	541.4	28.6	761	6	CD633904	CD633904
31	541.4	28.6	851	5	B0228581	B0228581
32	535.2	28.3	824	7	CK483217	CK483317
33	534	28.2	1003	5	B0056806	AGENCYCOURT
34	532.8	28.1	890	6	CA487493	AGENCYCOURT
35	531.8	28.1	535	4	BG086714	H3129F11-
36	530	28.0	1120	5	BM909505	AGENCYCOURT
37	529.2	28.0	760	4	BG753131	BM953131
38	528	27.9	1038	5	B0065633	AGENCYCOURT
39	526.2	27.8	1047	5	B0064127	AGENCYCOURT
40	524.8	27.7	865	4	BG172224	602333366
41	522.4	27.6	616	4	B1647932	603276014
42	521	27.5	818	7	CK483904	AGENCYCOURT
43	508	26.8	556	6	CA535696	CA535696
44	506.4	26.8	1044	5	B0065997	AGENCYCOURT
45	505.8	26.7	699	6	CD632908	55053832H

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION
AK004737	2899 bp	mRNA
AK004737	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013B2 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIA0537 (EC 2.7.1.-) [homo sapiens], full insert sequence.	linear HTC 03-APR-2004

ACCESSION	AK004737
VERSION	AK004737.2
KEYWORDS	GI:26334437
SOURCE	HTC; CAP trapper.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

REFERENCE
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

AUTHORS
Shimada, K., Itoh, M., Aizawa, K., Nagakoba, S., Sasaki, N., Carminci, P.,
Kunaba, R., Atiyama, Z., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsubara, S., Kawai, J.,
Okazaki, Y., Murakatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5

QY 1561 GAAGGCACTACCCCTAGACACCTTGGCTCCCTGAGCAACTGGCTCTCCATCTGCA 1620
1563 GAAGGCACTACCCCTAGACACCTTGGCTCCCTGAGCAACTGGCTCTCCATCTGCA 1712
QY 1621 GCCCGGCCGAGCCGCCCTCAGAGGCGCTGAGTGAAGACAGACATCTCTCCGAGTCC 1680
1713 GCCCGGCCGAGCCGCCCTCAGAGGCGCTGAGTGAAGACAGACATCTCTCCGAGTCC 1772
QY 1681 TTGACCAATTTGACTTGGCTGGAACGCTTCTCCGAAACCCCACTGAGGGGCTGATGCT 1740
1773 TTGACCAATTTGACTTGGCTGGAACGCTTCTCCGAAACCCCACTGAGGGGCTGATGCT 1832
QY 1741 GTGACCAACTGAGGGGCTTGAAGACACCTCTCCAGAGGCTGAAAGCTGAGTGGCAG 1800
1833 GTGACCAACTGAGGGGCTTGAAGACACCTCTCCAGAGGCTGAAAGCTGAGTGGCAG 1892
QY 1801 GAATCTTTGGGGGATGAGCTGCTTTCTCTGACAGACTGCGCAAGGAGTACTGACGCTTAC 1860
1893 GAATCTTTGGGGGATGAGCTGCTTTCTCTGACAGACTGCGCAAGGAGTACTGACGCTTAC 1952
QY 1861 AGACAGCCCTAGAGATCTGCTCAAGCTCAGC 1893
1953 AGACAGCCCTAGAGATCTGCTCAAGCTCAGC 1985
Db

RESULT 3
LOCUS AK034082 2869 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
cDNA library, clone:9330154N24 product:weakly similar to
PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)
[Homo sapiens], full insert sequence.
AK034082
VERSION AK034082.1 GI:26329672
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akizawa, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K.,
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2869)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://fantom.qualifiers

FEATURES
source
1. 2869
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9330154N24"
/db_xref="taxon:10090"
/clone="9330154N24"
/sex="male"
/tissue_type="dienecephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
98.2017
/note="unnamed protein product; putative
weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|060285,
evidence: FASTY, 56.8%ID, 98.6%length, match=1849)"
/codon_start=1
/protein_id="BAC28575.1"
/db_xref="GI:26329673"
/translation="MESGALLQRPQSPQASALASBAPSLADGLIKSPYLMKQAV
KRHHKHNLRHREPLETGLGTYGKAKRESGRVALKSRKDKIKQEDLIHR
RELEISGLNPHIILAIHEVGRSLVTPVENSILYIMVAYRQADLVYISSEPRLS
ERAPRFRFIVGSAIATYCHONGIVYHVDIKENILIDANKIKADRELSNLYKRGEL
QTGCSFLVSPFIVNKPVEVBSWISGLVLIIVHGTMPDGDHDTLVQISN
GASDFKPSADAGLIRMLMVNPTRAVLDEVAHVMWVGTTGVRGDBALREGH
PASDFKPSADAGLIRMSRSLNENGAFCFFQHYGVGSGTVGLRORHSLKRSRE
NDMAONIQGPADTSSRPKGSIKLPGIKKSSYSGSEVODPEOLRVPVTPQO
PVAVALSLPRKGIKTSRQESGIVYSPPRSEGLDADVPVSGPVPKQSPQASG
LIARKEIILKNGKFRYALRGTTPTFRSGLDOLASHPAPRBRPGAVSESDILS
ESFDQDLPERLPETPLRGCVSDNLKGLPQPSBGLKRWQBSLDSGCSLTDCQEV
TAAYRQALGICSKLS"

ORIGIN
Query Match 98.1%; Score 1857.4; DB 3; Length 2869;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
QY 1 ATGAGTGGTGGCTTACTTCAGCGCCGAGCGAGCTCCCTCGGCTCGGCTGACC 60

Db 98 ATGAGTTCGGGGGCTTACTTCAGCGCCGAGCCAGGCTCCCTGGGCTCCGCGCTTGGCC 157
Qy 61 TCGGAGAGCCCGCGCGCTGGCGGAGCGGCTCATCAAGTGGCTTAAACCTTGATGAAG 120
Db 158 TCGGAGAGCCCGCGCGCTGGCGGAGCGGCTCATCAAGTGGCTTAAACCTTGATGAAG 217
Qy 121 AAGGAGCGGTGAGAGCGGCGCATCAAGAAACAACAACCTGGCGGACCCCTAGAGATTCTG 180
Db 218 AAGGAGCGGTGAGAGCGGCGCATCAAGAAACAACAACCTGGCGGACCCCTAGAGATTCTG 277
Qy 181 GAGAGCTGGGCAAGGCGCATCAAGGAGGTGAAGAGGCAAGAGAGTCTGGGCGT 240
Db 278 GAGAGCTGGGCAAGGCGCATCAAGGAGGTGAAGAGGCAAGAGAGTCTGGGCGT 337
Qy 241 CTGGTGGCATCAAGTTCATCAGGAAGAAACAAATCAAGATGAGAGATCTGCTGAC 300
Db 338 CTGGTGGCATCAAGTTCATCAGGAAGAAACAAATCAAGATGAGAGATCTGCTGAC 397
Qy 301 ATTACGAGGAGATTTGATGATGATCTTCACTCAACCCCGCATCATTTGCCATGCAT 360
Db 398 ATTACGAGGAGATTTGATGATGATGATCTTCACTCAACCCCGCATCATTTGCCATGCAT 457
Qy 361 GAA-----GTGTTGAGATAGCAGCAAGATTGTGTTGTC 396
Db 458 GAACTGGGCAAGATCAGCTTAGTACGTGTGTTGAGATAGCAGCAAGATTGTGTTGTC 517
Qy 397 ATGAGATATCCAGCCGAGCGGATCTGTATGATTTACATCAGTGAACGCGCATGAT 456
Db 518 ATGAGATATCCAGCCGAGCGGATCTGTATGATTTACATCAGTGAACGCGCATGAT 577
Qy 457 GAGGAGGAGCGGCAAGGCTTTCTTCCGACAGATCGTGTCTGCGCTGCACTACTGCAACAG 516
Db 578 GAGGAGGAGCGGCAAGGCTTTCTTCCGACAGATCGTGTCTGCGCTGCACTACTGCAACAG 637
Qy 517 AACGAGATCGTTCCAGAGATCTCAAGTGGAAAAACCTCTTCAAGATGCAATGGAAC 576
Db 638 AACGAGATCGTTCCAGAGATCTCAAGTGGAAAAACCTCTTCAAGATGCAATGGAAC 697
Qy 577 ATCAAGATTCGTGACCTTTGGCTCTGCAACCTGTACCAACAAAGGCAAGTTCTTCAGAG 636
Db 698 ATCAAGATTCGTGACCTTTGGCTCTGCAACCTGTACCAACAAAGGCAAGTTCTTCAGAG 757
Qy 637 TTCTGTGGGAGCCCTCTCTAGCGCTCGCTGAGATAGTCAACGCGGAGCCCTATGTGGC 696
Db 758 TTCTGTGGGAGCCCTCTCTAGCGCTCGCTGAGATAGTCAACGCGGAGCCCTATGTGGC 817
Qy 697 CCAAGAGTGAAGCTGCTCTGAGGCTTCTCTGTACATCTGTGATGAGACCATG 756
Db 818 CCAAGAGTGAAGCTGCTCTGAGGCTTCTCTGTACATCTGTGATGAGACCATG 877
Qy 757 CCTTTGACGCGGAGATCAAAAACATGTTGAGAAATCACTGTAACGCGGCTTACCGT 816
Db 878 CCTTTGACGCGGAGATCAAAAACATGTTGAGAAATCACTGTAACGCGGCTTACCGT 937
Qy 817 GAGCCGCGCAAGCGCTCCGATGCTGTGGCTGTATCCGCTGATGATGTAACCC 876
Db 938 GAGCCGCGCAAGCGCTCCGATGCTGTGGCTGTATCCGCTGATGATGTAACCC 997
Qy 877 ACCGATGCGGCACTGAGAGATGTAAGCATTTGTGGGTCACTGGGGTTAAC 936
Db 998 ACCGATGCGGCACTGAGAGATGTAAGCATTTGTGGGTCACTGGGGTTAAC 1057
Qy 937 ACCGATGCGGCACTGAGAGCGCTCGTGAAGGTGGGCACTTATGTTGATTTGGC 996
Db 1058 ACCGATGCGGCACTGAGAGCGCTCGTGAAGGTGGGCACTTATGTTGATTTGGC 1117
Qy 997 CGGCGCTCTCATGCGGAGCTGTTAGTGTCTCTCGCGCCCTCTGAGAAATGAGCC 1056
Db 1118 CGGCGCTCTCATGCGGAGCTGTTAGTGTCTCTCGCGCCCTCTGAGAAATGAGCC 1177
Qy 1057 AAGGTGAGAGCTTCTTCAAGAGCAGAGTCCGGAAGTGAAGCACTGTACCTGGGCTG 1116

Db 1178 AAGGTGAGAGCTTCTTCAAGAGCAGACGTGCCGAGGTGAAGCACTGTACCTGGGCTG 1237
Qy 1117 GAGCGGCAATTTCTCTTAAAGAGTCCCGAAAGAAATGATGATGCTCAAAATCTGCA 1176
Db 1238 GAGCGGCAATTTCTCTTAAAGAGTCCCGAAAGAAATGATGATGCTCAAAATCTGCA 1297
Qy 1177 GGTGACCCGCTGAGATTCCTTCTGCGCCCTGGGAAGGAGCCTTAAGCTTCGAA 1236
Db 1298 GGTGACCCGCTGAGATTCCTTCTGCGCCCTGGGAAGGAGCCTTAAGCTTCGAA 1357
Qy 1237 GGCATTTCAAGAAAAAGTCTCTACCTCGTCAAGGAGGTACAGAGAGCCTTCAGAA 1296
Db 1358 GGCATTTCAAGAAAAAGTCTCTACCTCGTCAAGGAGGTACAGAGAGCCTTCAGAA 1417
Qy 1297 CTGACACGGTCTGATATCTCAAGGAGCCTGCTCTGCTGATCTCTGCTCCAGG 1356
Db 1418 CTGACACGGTCTGATATCTCAAGGAGCCTGCTCTGCTGATCTCTGCTCCAGG 1477
Qy 1357 AAAGGCACTCTTAAAGATCTGACAGCGGTGAATCTGTATTAACCTCTCGAGAGCC 1416
Db 1478 AAAGGCACTCTTAAAGATCTGACAGCGGTGAATCTGTATTAACCTCTCGAGAGCC 1537
Qy 1417 AGCGAGTCTGGGAACTCTTGAAGCAGATGATGTTGTGAGTGGGAGCCCGTGAAG 1476
Db 1538 AGCGAGTCTGGGAACTCTTGAAGCAGATGATGTTGTGAGTGGGAGCCCGTGAAG 1597
Qy 1477 CAGAAGTCTCAGAGCTTCAAGGCTCTCTTCACCGAGGCAATCTCAAACTCAAT 1536
Db 1598 CAGAAGTCTCAGAGCTTCAAGGCTCTCTTCACCGAGGCAATCTCAAACTCAAT 1657
Qy 1537 GCGAAGTCTCCGCAAGCCTTGAAGGACATACCCCTTAAGCACTTTGGCTCCCTGAG 1596
Db 1658 GCGAAGTCTCCGCAAGCCTTGAAGGACATACCCCTTAAGCACTTTGGCTCCCTGAG 1717
Qy 1597 CAAGTGGCTCTCTCCATCTCTGAGCCCGGAGCCCGCTCAAGGAGCTGTAGTGA 1656
Db 1718 CAAGTGGCTCTCTCCATCTCTGAGCCCGGAGCCCGCTCAAGGAGCTGTAGTGA 1777
Qy 1657 GACAGATCTGTCTCTCCAGTCTTGAACAAATGATGATGCTTGAACGCTTCCGAA 1716
Db 1778 GACAGATCTGTCTCTCCAGTCTTGAACAAATGATGATGCTTGAACGCTTCCGAA 1837
Qy 1717 ACCCACTGAGGGGCTGTGTCTGTGCAACCTGAAGGGGCTTGAAGCCTTCCCTCA 1776
Db 1838 ACCCACTGAGGGGCTGTGTCTGTGCAACCTGAAGGGGCTTGAAGCCTTCCCTCA 1897
Qy 1777 GAAGTCTGAAGGATGATGAGGAGATCCTTGGGAGTATGCTTTCTGACAGAC 1836
Db 1898 GAAGTCTGAAGGATGATGAGGAGATCCTTGGGAGTATGCTTTCTGACAGAC 1957
Qy 1837 TGCAGAGAGTGACTGACGCTTACAGACAAAGCCTTGAATGCTCAAGCTCAGC 1893
Db 1958 TGCAGAGAGTGACTGACGCTTACAGACAAAGCCTTGAATGCTCAAGCTCAGC 2014

RESULT 4
AY399041 1896 bp DNA linear GSS 12-DEC-2003
LOCUS AY399041
DEFINITION Mus musculus HCM0091 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY399041
VERSION AY399041.1 GI:39755030
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1896)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous

gene tr108
Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1896)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejalaval,A.,
Todd,M.A., Tanenbaum,D.M., Clivello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Stansky,D.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1896
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1896
/locus_tag="HCM0091"
ORIGIN
Query Match 93.7%; Score 1773; DB 9; Length 1896;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1773; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 1 ATGAGTGGTGGCTTACTTCAGCGCCGAGCGAGGCTCTCGGCTCCGCGCTGGCC 60
DB 1 ATGAGTGGTGGCTTACTTCAGCGCCGAGCGAGGCTCTCGGCTCCGCGCTGGCC 60
QY 61 TCGAGAGCGCCCGCGCGCTGGCGGAGGGCTCATCAAGTCGCTAAACCTGTATGAG 120
DB 61 TCGAGAGCGCCCGCGCGCTGGCGGAGGGCTCATCAAGTCGCTAAACCTGTATGAG 120
QY 121 AACGAGCGGTGAAAGCGGCAACATCAAAACAACCTCGGCGACCGCTTCAGATTCTG 180
DB 121 AACGAGCGGTGAAAGCGGCAACATCAAAACAACCTCGGCGACCGCTTCAGATTCTG 180
QY 121 AAGCAGCGGTGAAAGCGGCAACATCAAAACAACCTCGGCGACCGCTTCAGATTCTG 180
DB 121 AAGCAGCGGTGAAAGCGGCAACATCAAAACAACCTCGGCGACCGCTTCAGATTCTG 180
QY 181 GAGAGCTGGGGAAGGCACTTACGAGGAGTGAAGAGGCAAGAGAGCTGGGGCGT 240
DB 181 GAGAGCTGGGGAAGGCACTTACGAGGAGTGAAGAGGCAAGAGAGCTGGGGCGT 240
QY 241 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGACAGATTCTGTCTGAC 300
DB 241 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGACAGATTCTGTCTGAC 300
QY 301 ATACGAGGAGATTGATCATGTCTTCACTCAACCAACCCCATCATTTGCCATCAT 360
DB 301 ATACGAGGAGATTGATCATGTCTTCACTCAACCAACCCCATCATTTGCCATCAT 360
QY 361 GAAGTGTGAGATTGACAGAAATGATGATGTCATGAGATGACAGAGCGAT 420
DB 361 GAAGTGTGAGATTGACAGAAATGATGATGTCATGAGATGACAGAGCGAT 420
QY 421 CTGTATGATTACATCAGTGAAGCGGCAACGCTGAGTGAAGCGGCAAGGCAATTTCTT 480
DB 421 CTGTATGATTACATCAGTGAAGCGGCAACGCTGAGTGAAGCGGCAAGGCAATTTCTT 480
QY 481 CGACAGATCGTGTCTGCGCTGCACTACTGCGCAACGAGATGTTCAACCGAGATCTC 540
DB 481 CGACAGATCGTGTCTGCGCTGCACTACTGCGCAACGAGATGTTCAACCGAGATCTC 540
QY 541 AACCTGGAAGAACTCTCTTCTAGATGCAATGAAAGCACTCAAGATGCTGATCTTTGGCTC 600
DB 541 AACCTGGAAGAACTCTCTTCTAGATGCAATGAAAGCACTCAAGATGCTGATCTTTGGCTC 600
QY 601 TCCAACTGTACACAAAGGCAAGTCTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 660
DB 601 TCCAACTGTACACAAAGGCAAGTCTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 660
QY 661 TCGCTGAGATGATGCAACGAGAAACCTTATGTGGCCCAAGGTGACAGTGTCTCTG 720
DB 661 TCGCTGAGATGATGCAACGAGAAACCTTATGTGGCCCAAGGTGACAGTGTCTCTG 720
QY 720
DB 720

QY 721 GAGGTTCTCTGTACATCTGTGTGATGAGCAATGCGCTTTGACGGGCAAGATCATAA 780
DB 721 GAGGTTCTCTGTACATCTGTGTGATGAGCAATGCGCTTTGACGGGCAAGATCATAA 780
QY 781 ACACGTGTGAAGAAATCAATGATACGGGGCTTACCGTGAAGCGGCGCAAGCGTCCAGTCC 840
DB 781 ACACGTGTGAAGAAATCAATGATACGGGGCTTACCGTGAAGCGGCGCAAGCGTCCAGTCC 840
QY 841 TGTGACCTGATCCGGTGTGTTAATATGATGAAACCCCAACCGTGGGCACTAGAGAT 900
DB 841 TGTGACCTGATCCGGTGTGTTAATATGATGAAACCCCAACCGTGGGCACTAGAGAT 900
QY 901 GTAGCAGATCATTTGTGTGCTTAAATGATGAAACCCCAACCGTGGGCACTAGAGAT 960
DB 901 GTAGCAGATCATTTGTGTGCTTAAATGATGAAACCCCAACCGTGGGCACTAGAGAT 960
QY 961 CTGCGTGAAGGTGAGCAACCTGATGATGATCTTTGGCGGCGCTTCATAGCGGATCTGTTA 1020
DB 961 CTGCGTGAAGGTGAGCAACCTGATGATGATCTTTGGCGGCGCTTCATAGCGGATCTGTTA 1020
QY 1021 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGTGACAGCTTCAAGCAG 1080
DB 1021 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGTGACAGCTTCAAGCAG 1080
QY 1081 CACGTCCGGAGGTGAGCACTGTACCTTGGCTGAGCGGCAACATTTCTTAAAGAG 1140
DB 1081 CACGTCCGGAGGTGAGCACTGTACCTTGGCTGAGCGGCAACATTTCTTAAAGAG 1140
QY 1141 TCCGAAAGGAATGACATGCTCAAAATCTGAAAGTGAACCGGCTGAGATCTCT 1200
DB 1141 TCCGAAAGGAATGACATGCTCAAAATCTGAAAGTGAACCGGCTGAGATCTCT 1200
QY 1201 TCTGGCTCTGGCAAGAGCGCTTAAAGTTCGAAAGGCAATTCGAAAGAAAGTCTCT 1260
DB 1201 TCTGGCTCTGGCAAGAGCGCTTAAAGTTCGAAAGGCAATTCGAAAGAAAGTCTCT 1260
QY 1261 ACCTGTCAGGGAGGTATCAGAGAGACCTCAGAACTCAGACCGGCTGTATCTCA 1320
DB 1261 ACCTGTCAGGGAGGTATCAGAGAGACCTCAGAACTCAGACCGGCTGTATCTCA 1320
QY 1321 GGGCAGCTGTCTCTGTGTATCTCTGCTCCAGAGAAAGCATCTTAAAGAGTCTGCA 1380
DB 1321 GGGCAGCTGTCTCTGTGTATCTCTGCTCCAGAGAAAGCATCTTAAAGAGTCTGCA 1380
QY 1381 CAGCGTGAATCTGTATCTCTCTCCAGAGCGGATCTGGGGAACCTTAAAG 1440
DB 1381 CAGCGTGAATCTGTATCTCTCTCCAGAGCGGATCTGGGGAACCTTAAAG 1440
QY 1441 GCCAGTATGTGTTTGTGATGAGGAGACCCGCTGAGAGCAAGTCTCCAGAGCTTCAAGG 1500
DB 1441 GCCAGTATGTGTTTGTGATGAGGAGACCCGCTGAGAGCAAGTCTCCAGAGCTTCAAGG 1500
QY 1501 CTCCTCTCCACCGCAAGGCAATCTCAATGCAAGTCTTCCGACAGCTTAA 1560
DB 1501 CTCCTCTCCACCGCAAGGCAATCTCAATGCAAGTCTTCCGACAGCTTAA 1560
QY 1561 GAAAGCACTACCCCTGACACTTGGCTCCCTGACCAATGAGCTCTCTCCATCTGCA 1620
DB 1561 GAAAGCACTACCCCTGACACTTGGCTCCCTGACCAATGAGCTCTCTCCATCTGCA 1620
QY 1621 GCCCGGCGCAGCGCCCTCAGGGGCTGATGAGGACAGAGATCTCTCCAGATCC 1680
DB 1621 GCCCGGCGCAGCGCCCTCAGGGGCTGATGAGGACAGAGATCTCTCCAGATCC 1680
QY 1681 TTTGACCAATTTGACTTCTGTAAGCTTTCCGAAACCCCACTGAGGGCTGTGTGCT 1740
DB 1681 TTTGACCAATTTGACTTCTGTAAGCTTTCCGAAACCCCACTGAGGGCTGTGTGCT 1740
QY 1741 GTGAGCAACTGAGGGGCTTGAAGAGCTCTCCAGAGTGTGAAGGATGAGTGTGAG 1800
DB 1741 GTGAGCAACTGAGGGGCTTGAAGAGCTCTCCAGAGTGTGAAGGATGAGTGTGAG 1800

QY	1861	AGACAAAGCCCTAGAGATCTGCTCAAAAGCTCAGC	1893
QY	1801	GAATCCCTGGGGGAGTAGCTGCTTTTCTCTGACAAGCTCCCAAGAGAGTACATCGACCCAC	1868
Db	1801	GAATCTTTGGGGGAGTAGCTGCTTTTCTCTGACAAGCTCCCAAGAGTATCATCGACCTAC	1868
QY	1861	AGACAAAGCCCTAGAGATCTGCTCAAAAGCTCAGC	1893
Db	1861	AGACAAAGCCCTAGAGATCTGCTCAAAAGCTCAGC	1893

	RESULT 5	
HSM801859		
LOCUS		
HSM801859	3443 bp	mRNA
Homo sapiens		linear
DEFINITION		HTC 22-SEP-2004
ACCESSION		
ATCC6037		
		CDNA DKEFZp334J037 (From clone DKEFZp334J037).

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE

100

TITL

COMMENT

Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center) at the Heinrich-Heine-University, Dueseldorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZ3434037) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:

<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZ3434037>

Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna>.

FEATURES
source

```

/gene="Homo sapiens"
/mol_type="mRNA"
/db_xref="RDP:DKFzp43J037"
/db_xref="taxon:9606"
/clone="DKFzp43J037"
/tissue_type="testis"
/clone_idb="434 (synonym: hne3) . Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein"
1..3443
/gene="DKFzp43J037"
128..2014

```

[illegible]

Query Match	71.6%;	Score 1355.4;	DB 3;	Length 3443;
Best Local Similarity	84.0%;	Pred. No. 0;		
Matches 1562; Conservative	0;	Mismatches 276;	Indels 21;	Gaps 2;

QY	47	CCCTCCGCGCTGGCTCTGCGGAGAGCGCCCTGCGCTGGCTCGACCGGCTCATCAAGTCGTCTA	106
Db	162	CCACTCCCTGGGCGCGCAGAGCTAGCCCGGCGCGCTGGCGGAAAGGGCTGTATCAATTCGCGCA	221
QY	107	AACCTGTGATGAAGAAAGCAGCGCGGTGAAGGCGGACCATCAACAACAACCTTCGCGCACCC	166
Db	222	AGCCCCCTTAAGAAAGAAAGCAGGCGGTGAAGGCGGACCAACAAGCAACATCTCGCGCACCC	281
QY	167	GCTACGAGTTCCTGAGAGCGCTGGGCAAGGGCACTTACCGGGAAAGTGAAACAAGGACACAG	226
Db	282	GCTACGAGTTCCTGAGAGCCCTTGGCGAAAGGCACTTACCGGGAAAGTGAAAGAGGCGCGG	341
QY	227	AGAGCTTCGGGCGCTGTGTGGCCATCAAGTCCATCAGAAAGAACAATTCAAAGATGAGC	286
Db	342	AGAGCTTCGGGCGCGCTGTGTGGCCATCAAGTCCATCCGAAAGACAATAATCAAGATGAGC	401
QY	287	AGGATCTGCTGCACTATACGGAAGGGAATTGAGTTCATGTCTTCACTCAACCAACCCACACA	346
Db	402	AAGATCTGATGACATACGAGGGAAGTTGAGATCATGTATCATCTCAACCAACCTTCACA	461
QY	347	TCATTCGCATCCATGGAAGTGTGTGAAGATAGCAGCAAGATTGTGATGTTCATGAGATATG	406
Db	462	TCATTCGCATCCATGGAAGTGTGTGAAGACAGCAAGATCGTGAATGTATGAGATATG	521
QY	407	CCAGCCGAGGCGATCTGTATGATTACATCAGTAGAGCGGCCACGGCTGAGTGAAGCGGACG	466
Db	522	CCAGCCGAGGCGACCTTTATGACTATCATCAGCAGAGCGGCGAGCTCAGTAGAGCGGAGAG	581
QY	467	CCAGGCAATTTCTTCGACAGATGTCGTCTGCGCCCTGCACTATCGGCCACAGAAAGGGATG	526
Db	582	CTAGGCAATTTCTTCGCGCAGATGTCTCTGCGCGTGACTATATGCCATTCAGAAAGAGTTG	641
QY	527	TTTACCGAGATCTCAAGCTGGAAGAAATCATCTTTAGATGCGCAATGGAACATCAAGATTG	586
Db	642	TTCAACCGAATCTCAAGCTGGAAGAAATCTCTTGAGATGCGCAATGGAAGATTCAGATTG	701
QY	587	CTGACTTTGGCGCTCTCCAACTGTACCAACAAGGCAAGTTCTCCAGACGTTCTGTGGGA	646
Db	702	CTGACTTCGGCGCTCTCCAACTGTACCACTCACTCAACAAGGCAAGTTCTCGCAGACATTCGTGGGA	761
QY	647	GGCCTCTTACGCGCTGCGCTGAGATATGTCAACGGGAAGCCCTTATGTGGCGCCAGAGTGG	706
Db	762	GGCCCTCTATGCTCTGCGCAAGATTTGTCAATGGGAAGCCCTTACACAAGGCCCAAGGTGG	821
QY	707	ACAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGACATGCGACACATGCCCTTTGACG	766
Db	822	ACAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGACATGCGACACATGCCCTTTGATG	881
QY	767	GGCAGATCATAAACAATGTGTGAGCAAAATCAGTACGGGGCTTACCTGTAGCCGCGCA	826
Db	882	GGCATATCAATTAAGATCTTAGTGAAACAAGATCAGCAACGGGGCGCTTACCGGGAGCCACTTA	941
QY	827	AGCGGTGCCATGCTGTGTGCGCTGTATCCGGGTGGCTGTAAATGTGGAACCCCAACCGGTGCGG	886
Db	942	AACCTCTGTATGTCTGTGGCTGTATCCGGGTGGCTGTGTATGTGGAACCCCAACCGGTGCGG	1001
QY	887	CCACACTGAGAGATGTATGACAGTCAATGTGTGGGTCAACTGGGGTTTACACACCGGAGTCG	946
Db	1002	CCACCTGTGAGATGTGTGCAAGTCACTGGGTGGTCAACTGGGGCTTACGCCAACCGAGTGG	1061
QY	947	GGGAAACAAGAGCCCTTGCGGTGAGGGTGGGACCTTAGTGTGACTTTTGCGCGGCGCTCCA	1006
Db	1062	GAGAGCAGAGAGCTCCGCAATGAGGGTGGGACCTCTGGCAGTGACTCTGCGCGGCTCCA	1121
QY	1007	TGGCGGACTGTGTAAAGTCGTCTCTCGCGCCCTCTCTGTGAGAAATGAGCGCAAGTGTGCA	1066
Db	1122	TGGCTATCTGGCTCCGCGCTTCTCTCCGCGCCCTCTCTGTGAGAAATGAGCGCAAGTGTGCA	1181

QY 1067 GCTTCTCAAGACGACGTCGCGGAGGTGAGACCTGTAACCTGGGCTGAGCGGCAAC 1126
 Db 1182 GCTTCTCAAGACGACGTCGTCGCGGAGGACCAACCCCTGGCTGAGCGGCAAC 1241
 QY 1127 ATTCTCTTAAGAGTCCCGAAGAGATGACATGGCTCAAAATCTGCAAGTACCCGG 1186
 Db 1242 ATTCTCTTAAGAGTCCCGAAGAGATGACATGGCTCAAAATCTGCAAGTACCCGG 1301
 QY 1187 CTGAGATACCTCTTCTGCGCCCTGGCAAGACGCTTAAGCTTCCGAAAGGCAATTCCTCA 1246
 Db 1302 CTGATGACATGCGCCCATGCGCTGGCAAGACGCTTCAAGCTTCCGAAAGGCAATTCCTCA 1361
 QY 1247 AGAAAAAGTCTCTTCTGCTGCTGAGGAGGTACAGAGAACCTTCAGAACTCAGACCGG 1306
 Db 1362 AGAAGAGAGTGTGAGCTCTGCAAGAGGGGTACAGAGAACCTTCAGAACTCAGACCGG 1421
 QY 1307 TGCTGTATCTCCAGGGGACGCTTCCCTGCTGATCCCTGCTCCCAAGAAAGGCAATCC 1366
 Db 1422 TCCCTGTGAGCCCGGAGGAGGCTGCCCC-----CGCTGCTCCCAAGAGGCAATTC 1472
 QY 1367 TTAAGAGTCTGCAAGGAGTGAATCTGTTACTACTCTCTCCAGAGCCGAGGAGTCTG 1426
 Db 1473 TCAAGAGCCCGGAGGAGGAGTGTGCTACTACTCTCTCCAGAGCCGAGGAGTCTG 1532
 QY 1427 GGAAGTCTTGAAGCGCAGTGAATGTTTGTGAGTGGGAGCCCGTGGAGGAGGAGTCTG 1486
 Db 1533 GGAAGTCTTGAAGCGCAGTGAATGTTTGTGAGTGGGAGCCCGTGGAGGAGGAGTCTG 1592
 QY 1487 CAGAGGCTTCAAGGCTCTCTCTCAAGCGCAAGGAGTCTCAAACTCAATGCGCAATGCT 1546
 Db 1593 CAGAGGCTTCAAGGCTCTCTCTCAAGCGCAAGGAGTCTCAAACTCAATGCGCAATGCT 1652
 QY 1547 CCGGACAGGCTTGAAGAGGACATACCCCTGACACCTTGGCTCCCTGAGCAACATGGCT 1606
 Db 1653 CCGGACAGGCTTGAAGAGGACATACCCCTGACACCTTGGCTCCCTGAGCAACATGGCT 1712
 QY 1607 CCTCCCATCTGACGCGCGCCGACCGCCCTCAAGGAGTGTGAGTGAAGAGGACATCC 1666
 Db 1713 CACCTGCGCCCTGAGCGCGCGCCGACCGCCCTCAAGGAGTGTGAGTGAAGAGGACATCC 1772
 QY 1667 TGTCCTCGAGTCTTGAACCAATTGACCTTGGCTGAAAGCTTTCGGAACCCCACTGA 1726
 Db 1773 TGTCCTCGAGTCTTGAACCAATTGACCTTGGCTGAAAGCTTTCGGAACCCCACTGA 1832
 QY 1727 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGGAGAGGCTCCCTCAAGAG----- 1780
 Db 1833 GGGGCTGTGTGTGTGAGCAACCTGAGGGGCTTGGAGAGGCTCCCTCAAGAGGGCCCTG 1892
 QY 1781 -----GTCTGAAGCGATGTGAGGAGATCTTGGGGATAGCTGCTTTCTCTGACAG 1834
 Db 1893 GAAGCTGCTGAGGCGCTGGCGGAGATCTTGGGGAGACGCTGCTTTTCCCTGACAG 1952
 QY 1835 ACTGCGAAGAGTGAAGTGAAGCTTCAAGCAAGGCTTGAAGATGTGCTCAAGCTCAGC 1893
 Db 1953 ACTGCGAAGAGTGAAGTGAAGCTTCAAGCAAGGCTTGAAGATGTGCTCAAGCTCAGC 2011
 RESULT 6
 AY399039
 LOCUS 1887 bp DNA linear GSS 12-DEC-2003
 DEFINITION Homo sapiens HCM0091 gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY399039
 VERSION AY399039.1 GI:39755028
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1887)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

TITLE Adams,M.D. and Cargill,M.
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14673302
 REFERENCE 2 (bases 1 to 1887)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 location/Qualifiers
 1..1887
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..1887
 /locus_tag="HCM0091"
 ORIGIN
 Query Match 66.6%; Score 1261; DB 9; Length 1887;
 Best Local Similarity 78.4%; Pred. No. 0;
 Matches 1458; Conservative 0; Mismatches 380; Indels 21; Gaps 2;
 QY 47 CTTCCGCGCTGGGCTCGGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAATGCGCTTA 106
 Db 35 CCACTCCCTGGCGCGAGAGCTAGCCCGCGCTGGCGGAGGCGCTCATCAATGCGCGCA 94
 QY 107 AACCTGAGAGAGAGACAGCGGCTGAGAGGCGGACCATCAACACAACTCGGCGAC 166
 Db 95 AGCCCTTAATGAAGAGAGAGAGGCTGAGAGGCGGACCAACACAACTCGGCGAC 154
 QY 167 GCTACGAGTCTCTGAGACGCTGGGCAAGGACCTTACGAGGAGTGAAGAGGACAGAG 226
 Db 155 GCTACGAGTCTCTGAGACGCTGGGCAAGGACCTTACGAGGAGTGAAGAGGCGG 214
 QY 227 AGAGCTGGGCGCTGTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 286
 Db 215 AGAGCTGGGCGCTGTGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGC 274
 QY 287 AGAGTCTGCTGACATACGAGAGGAGTGAAGTCAATGCTTCACTCAACCCCA 346
 Db 275 AAGATCTGAATGACATACGAGAGGAGTGAAGTCAATGCTTCACTCAACCCCA 334
 QY 347 TCATTGCGATTCATGAAGTGTGAGAAATAGACAGCAAGATTGTGATTGTGAGATAG 406
 Db 335 TCATTGCGATTCATGAAGTGTGAGAAATAGACAGCAAGATTGTGATTGTGAGATAG 394
 QY 407 CCAAGCGAGGCGATCTGATGATTACATCAAGTGAAGCGGCAAGCTGATGAGCGGAGC 466
 Db 395 CCAAGCGAGGCGATCTGATGATTACATCAAGTGAAGCGGCAAGCTGATGAGCGGAGC 454
 QY 467 CCAAGCATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCAACAGAGGATG 526
 Db 455 CTAGGCAATTTCTTCCGACAGATGCTGTGCTGCTGCACTACTGCAACAGAGGATG 514
 QY 527 TTCAACGAGATCTCAAGCTGGAAGACATCTTCTGATGAGCAATGAGAACTCAAGATTG 586
 Db 515 TCACGAGATCTCAAGCTGGAAGACATCTTCTGATGAGCAATGAGAACTCAAGATTG 574
 QY 587 CTGACTTGGCGCTCTCAACCTGTACCAAGGCAAGGTTCTTCAGAGCTTCTGTGGGA 646
 Db 575 NNN 634
 QY 647 GCCCTCTCTAGCGCTGCGCTGAGATGATCAACGAGGAGCCCTATGTGGGCCAGAGTGG 706
 Db 635 NNN 694
 QY 707 ACAAGTGTCTCTGGGCGTCTCCTGTACATCTGTGTGATGACCATGCGCTTTGACG 766

Db	695	ACAGCTGGATCCCTGGGTGTTCTCCCTCTACATCTCGTGCAATGGACACATGCCCTTTATGT	754
OY	767	GGCAGGATCATATAAACACTGGTGAAAGCAATCATGAACGGGGCTTACCGTGAAGCCGCCA	826
Db	755	GGCATGACCATTAAGATCCCTAGTAAACAAGATCAAGACGGGGCTTACCGGAGGCCACTTA	814
OY	827	AGCGTCCGATAGCCGTGAGCCCTGATACCGGTGCTCTGTAAATGTGTAAACCCCAACCCGTCCGG	886
Db	815	AACCTCTGATGCTCTGTGGCCCTGATCCGGTGGCTGTGATGTTGAACCCCAACCCCGCGGG	874
OY	887	CCACACTGGAGGATGTATGACCAAGTCAATTGGTGGGTCAACTGGGGTTATACACACCGGAGTGG	946
Db	875	CCACCTTGAGAGGATGTGGCCAGTCACTGGTGGTCAACTGGGGCTTACCGCACCGGAGTGG	934
OY	947	GGGAAACAGAAAGCCCTGCGTGAAGGGTGGGCAACCTTAGTGTGTACTTTGGCCCGGGCTTCA	1006
Db	935	GAGAGCAAGAGGCTCCGATGATGAGGGTGGGCAACCTGAGCATGTACTGCCCCGGCTTCA	994
OY	1007	TGGGGGATTTGGTATAGTGGCTCTCTGGCCGCCCTCTCTGTGAAGATTTGAACTCAAGTGTGTCA	1066
Db	995	TGGCTGATCTGGCTCGGGGCTTCTCCGCCCTCTCTGTGAAGATTTGAGGCTCAAGGTTGTCA	1054
OY	1067	GCTTCTTCAAGCAGCAGCTGCGCGGAGGTGGAAGCACTGTACTGTGGCTGGAGCGGCAAC	1126
Db	1055	GCTTCTTCAAGCAGCATGACACTGTGTGGGGGAAAGCACACCTTGCTGGAGCGCCAGC	1114
OY	1127	ATTCTCTTAAAGAGTCCGGAAGGAAATGACATGGCTCAAAATCTGCAGAGTGAACCGG	1186
Db	1115	ATTGCTCAAGAAAGTCCGCAAGGAAATGACATGGCCCAATCTCCACAGTGAACGG	1174
OY	1187	CTGAGGAATCACTCTTCTGCGCTCTGGCAAGACAGCTTAACTTTCGAAAGGACTTCTCA	1246
Db	1175	CTGATGACATCGCCCATGCGCTTGGCAAGAGCACTCAAGGTGCAAAAGGGCAATTCTCA	1234
OY	1247	AGAAAGATCCCTTACTCTGTCAGGGGAGGTACAGAGAGACCTCAGGAATCAGACCGG	1306
Db	1235	AGAAAGATGTCAAGCTCTGTGCAAGAGGGGTACAGAGAGACCTTCCGAGCTCAAGCCAA	1294
OY	1307	TGCTGATATCTCAGAGGAGCGCTGTCCCTGTGTATCCCTGTCTCCCAAGAAAGCATCC	1366
Db	1295	TCCCTGCGAGCCAGGGGCAAGGCTGCC-----CCGTGTGCCAAGAAAGGCATTC	1345
OY	1367	TTAAGAGATCTCGACAGCTGATATTTGGTACTACTCTCTCCAGAGCCACGAGTGTG	1426
Db	1346	TCAAGAAACCCCGACAGCGCAAGTCTGTGCTACTACTCTCTCCGAGCCCAAGTGAATCTG	1405
OY	1427	GGGAATCTTATGAGCGCGCAGATGATGTTTGTGAGTGGGAGCCCGTGGAGCAGAAATCTC	1486
Db	1406	GGAGACTTGTGAGCGCAGGGCAACGTTTGTGTGAGTGGGAAATCCCAAGAGCAAGACCTTC	1465
OY	1487	CACAGGCTTACAGGGCTCTCTCTCCACCGCAAGGGCAATTCTCAAACTCAATGGCAAGTCT	1546
Db	1466	CGAAGGCTTACAGGGGTGCTCTCTCATTCGCAAAAGCATCTCTCAAACTCAATGGCAAGTCT	1525
OY	1547	CCGCGACACCTTATGAAAGGCACTACCCCTAGACACTTTGGCTCTCCCTGAGCCAATCGGCT	1606
Db	1526	CCGAGACACGCTTGAAGCTCGCGGGCCCAACCACTTGGGCTCTCTGATGAATCGGCTC	1585
OY	1607	CCCTCCATCTCGACGCCGGCGCCAGCGCCCTCAGAGGGCTGTGAGTGAAGACACAGATCC	1666
Db	1586	CACCTCGCCCCCTGGCCCGGGCCAGGCCAACCTCTCAGGGGCTGTGAGCAGAGCACAGATCC	1645
OY	1667	TGCTCTCCGAGATCTTTGTAACCAATTGACTTGGCTGGAACGTCTTCCGAAACCCCACTGA	1726
Db	1646	TGCTCTCTGAGATCTTTTGAACAGAGCTTGAACCTTGTCTGAACGGCTCCCAAGAGCCCCCACTGC	1705
OY	1727	GGGGCTGTGTGTCTGTGTGCAACCTGAAGGGGCTTGAAGCAGCTTCTCCACAGAG-----	1780
Db	1706	GGGGCTGTGTGTCTGTGTGCAACCTCAAGGGGCTTGAAGAGGCCCTCCCTCAAGAGGGCCCTG	1765
OY	1781	-----GTCTGAAGCATGTGTGTGGCAGGAATCTTGGGGGATAGCTGTCTTTCTGTGACAG	1834

Db	1766	GAAGTGTCCCTGAGGCGCTGGCGGCGAGATTCCTTTGGGGGACAGCTGCTTTTCCCTGACAG	1823
Qy	1835	ACTGCCCAAGAGGTGACTGCAAGCCTTACAGACAGACCCCTAGAAATGCTGCTCAAGCTCAGC	1893
Db	1826	ACTGCCCAAGAGGTGACAGCAGCCTTACCGACAGGCACTGAGGGTCTGCTCAAGCTCAGC	1884
RESULT 7			
LOCUS	AY399040		
DEFINITION	Pan troglodytes HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence.		
ACCESSION	AY399040		
VERSION	AY399040.1		
KEYWORDS	GI:39755029		
SOURCE	GSS.		
ORGANISM	Pan troglodytes (chimpanzee)		
REFERENCE	Pan troglodytes		
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
TITLE	1 (bases 1 to 1887)		
JOURNAL	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,		
PUBMED	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,		
REFERENCE	Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,D.J.,		
AUTHORS	Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous		
JOURNAL	gene tios		
PUBMED	Science 302 (5652), 1960-1963 (2003)		
REFERENCE	2 (bases 1 to 1887)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,		
TITLE	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,		
JOURNAL	Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,D.J.,		
PUBMED	Adams,M.D. and Cargill,M.		
REFERENCE	Direct Submission		
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,		
JOURNAL	Rockville, MD 20850, USA		
PUBMED	This sequence was made by sequencing genomic exons and ordering		
REFERENCE	them based on alignment.		
FEATURES	Location/Qualifiers		
SOURCE	1.1887		
gene	/organism="Pan troglodytes"		
ORIGIN	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..>1887		
	/locus_tag="HCM0091"		
Query Match	50.1%;	Score 948.4;	DB 9;
Best Local Similarity	60.0%;	Pred. No. 2.2e-236;	
Matches 1116;	Conservative 0;	Mismatches 722;	Indels 21;
			Gaps 2;
Qy	47	CCTCCGCGCTGAGCGCTGAGAGCGCCCGCGCGTGGGGGAGCGGGCTCATCAAGTCGCTTA	106
Db	35	CCACTCTCCCTGGCGCGCAGACTAGCCCGCGCTGGGGAGAGGGCTGATCAAGTCGCCCA	94
Qy	107	AACCTCTGATGAGGAAGACAGCGCGGTGAGCGGACCATCAACAACCAACTCGCGCACCC	166
Db	95	AGCCCCCTAATGAAGAAGACAGCGCGGTGAGCGGACCATCAACAAGCAACACTCGCGCACCC	154
Qy	167	GCTACGAGTTCTGAGAGCGCTGGGCAAGGGCACTTACGGGAAGGTGAAGAAGGCAAG	226
Db	155	GCTACGAGTTCTGAGAGCGCTGGGCAAGGGCACTTACGGGAAGGTGAAGAAGGCGCGG	214
Qy	227	AGAGCTCGGGGCGCTGAGGCGCATCAAGTCATCAAGGAAGAACAATCAAGAGTAGC	286
Db	215	AGAGCTCGGGGCGCTGAGGCGCATCAAGTCATCAAGGAAGAACAATCAAGAGTAGC	274
Qy	287	AGAGCTCGGCGCATCAAGGAGGAGATTTGAGATCATGTCTTCACTCAACACCCGACCA	346
Db	275	AAGATCTGATGCACTAAGGAGGAGATTTGAGATCATGTCTTCACTCAACACCCGACCA	334
Qy	347	TCAATGTCATTCATGAGAGTGTGAGAAATGACGAAGAATTTGATTGATCGAGATATG	406

[illegible]

OY		1487	CACAGGCTTACGGGCTCCTCCTCCATCCAGCGAAAGGCATTTCCAACTCAATGGCAAGTTCT	1546
Dd		1466	NN	1525
OY		1547	CCCGCACAGGCTTATGAAGAAGTAACCCTTAGACACTTTGGCTCCTTGACCAATCGGCTT	1606
Dd		1526	NNNNGACAGGCTTGAGACTGGCGGCCCCCGAACACACTTGGCTCCTTGATGAACTCGCCC	1585
OY		1607	CCTCCCATCTGTGACAGCCCCGACGCTGCCCTTACAGGGGCTGTGAGTGAAGACAGCATCC	1666
Dd		1586	CATCTCGCCCCCTGGCCCCGGCGACGCACTTCAGGGGCTGTGAAGGAGAGACAGCATCC	1645
OY		1667	TGTCCTCGAGTCTTTGACCAATTTGAGACTTGCCTGAAGCTCTCCGAAACCCCACTGA	1726
Dd		1646	TGTCCTCGAGTCTTTGACCAAGTGTGACTTGTGCTGAACCGCTCCAGAGCCCCACTGC	1705
OY		1727	GGGGCTGTGTCTGTGGACAACCTGAGGGGGCTTGAAGACCTTCCTCTGAAAG-----	1780
Dd		1706	GGGGCTGTGTGTGTGTGGAACAACCTCATATGGGGGCTTGAAGAGCCGCCCTCTGAGAGGCCCTG	1765
OY		1781	-----GTCTGAACCATGTGTGGAGAAATCTTTGGGGGATAGTGTCTTTCTGTGACAG	1834
Dd		1766	GAACTGCTGTGAGCGCTGGCGGCGCAGAGATCTTTGGGGGACAGCTGTCTTTCCCTGACAG	1825
OY		1835	ACTGCGAAGAGGTACTGACGACGCTGACAGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC	1893
Dd		1826	ACTGCGAAGAGGTACACGACGACCTTACCGGACGACACTGAGGGGTCTGCTAAAGCTCAGC	1884
RESULT 8				
LOCUS	CR749209	7004 bp	mRNA	linear HTC 19-AUG-2004
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686F01113 (from clone DKFZp686F01113).			
ACCESSION	CR749209			
VERSION	CR749209.1	GI:51476151		
KEYWORDS	HTC.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
REFERENCE	Bukatyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 7004) Ostenwaelder,B., Obermaier,B., Deutschenbaaur,S., Schnapp,A., Mewes,H.W., Weill,B., Amid,C., Oeangner,A., Fobo,G., Han,M. and Wiemann.S.			
AUTHORS	The German CDNA Consortium			
COMMENT	Direct Submission Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuheberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-helidelberg.de; Sequenced by Medigenomix (Martinsried/Germany) within the CDNA Banking consortium of the German Genome Project. This clone (DKFZp686F01113) is available at the RZPD Deutsches Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686F01113 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .			
FEATURES				
SOURCE	location/Qualifiers			
	1..7004			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="DKFZp686F01113"			
	/issue_type="Fetal kidney"			
	/clone_id="686 (synonym: hlcc3). Vector psportl_sfi, host			
	DH10B; sites SfiI + SfiIB"			
	/dev_stage="Fetal"			
	/note="hypothetical protein, N-terminus truncated, not			
	fully spliced"			
	1..7004			
gene	/gene="DKFZp686F01113"			

/db_xref="taxon:10090"
 /clone="IMAGE:5718428"
 /tissue="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH_BMAP_F10"
 /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an end
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is TGAGAGACC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 36.1%; Score 683.6; DB 6; Length 726;
 Best Local Similarity 98.8%; Pred. No. 3e-167;
 Matches 720; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

```

QY 800 GTAAACGGGGCTTAC-CGTGAGCCGCCAAGCCGTCGATGCTGCGCTGATCCGGTGG 858
      |||||||
DB 1 GTAAACGGGGCTTACNCGTAGCCGCCCAAGCCGTCGATGCTGCGCTGATCCGGTGG 60
      |||||||
QY 859 CTGTTAATGATGAACCCCAACCCGTCGGGCCCACTGAGAGATGTAGCACTGATGTGG 918
      |||||||
DB 61 CTGTTAATGATGAACCCCAACCCGTCGGGCCCACTGAGAGATGTAGCACTGATGTGG 120
      |||||||
QY 919 GTCAACTGGGGGTTACACACCGGAGTGGGGGAAAGAGAGCCCTGCTGAGAGGTGGGAC 978
      |||||||
DB 121 GTCAACTGGGGGTTACACACCGGAGTGGGGGAAAGAGAGCCCTGCTGAGAGGTGGGAC 180
      |||||||
QY 979 CCAAGAGGTGACTTTGGCCCGGCTCCATGCGGACACTGCTGCTGCTGCTGCTGCTGCTG 1038
      |||||||
DB 181 CCAAGAGGTGACTTTGGCCCGGCTCCATGCGGACACTGCTGCTGCTGCTGCTGCTGCTG 240
      |||||||
QY 1039 CTCCTGAGATGAGAGCAAGTGTGAGCTTCTTCAAGAGCACTGCGCGGAGGTGGA 1098
      |||||||
DB 241 CTCCTGAGATGAGAGCAAGTGTGAGCTTCTTCAAGAGCACTGCGCGGAGGTGGA 300
      |||||||
QY 1099 AGCACTGACCTGGGCTGAGAGCCGCAATCTTCTTAAAGATCCCGAAAGAGATGAC 1158
      |||||||
DB 301 AGCACTGACCTGGGCTGAGAGCCGCAATCTTCTTAAAGATCCCGAAAGAGATGAC 360
      |||||||
QY 1159 ATGGCTAAATATGCAAGTGTGAGCTGCGGCTGAGATCTTTTCCGCTGCGAAAGAC 1218
      |||||||
DB 361 ATGGCTAAATATGCAAGTGTGAGCTGCGGCTGAGATCTTTTCCGCTGCGAAAGAC 420
      |||||||
QY 1219 AGCCTTAAGCTTCCGAAGGATCTCAAGAAAGATCTTACTGCTGCGGAGAGATA 1278
      |||||||
DB 421 AGCCTTAAGCTTCCGAAGGATCTCAAGAAAGATCTTACTGCTGCGGAGAGATA 480
      |||||||
QY 1279 CAGAGAGACCTTCAGAACTCAGACCGGTGCTGATCTCCAGGGAGACCTGCTCCCTGCT 1338
      |||||||
DB 481 CAGAGAGACCTTCAGAACTCAGACCGGTGCTGATCTCCAGGGAGACCTGCTCCCTGCT 540
      |||||||
QY 1339 GTATCCTGCTCCCAAGAAAGGATCTTAAAGATCTGCAAGGCTGAATCTGCTTAC 1398
      |||||||
DB 541 GTATCCTGCTCCCAAGAAAGGATCTTAAAGATCTGCAAGGCTGAATCTGCTTAC 600
      |||||||
QY 1399 TACTCCTCTCAGAGCCAGAGATCTGGGAACTTAAAGAGCAGATGATGTGTTGTG 1458
      |||||||
DB 601 TACTCCTCTCAGAGCCAGAGATCTGGGAACTTAAAGAGCAGATGATGTGTTGTG 660
      |||||||
QY 1459 AGTGGGAGCCCGTGAAGAGAGTCTTCAAGGCTTCAAGGCTCTTCTTCAAGGAG 1518
      |||||||

```

DB 661 AGCGAGNCA-CCCGTGAAGCAAG--TCTCAGAGCTTCAAGGCTCTCTCCACCGCAAG 717
 QY 1519 GGCATTCTC 1527
 |||||||
 DB 718 GGCATTCTC 726

RESULT 10
 CO424322 672 bp mRNA linear EST 06-JUL-2004
 LOCUS U1-M-HUO-cory-a-14-0-U1.r1 NIH_BMAP_HUO Mus musculus cDNA clone
 DEFINITION IMAGE:30665389 5', mRNA sequence.
 ACCESSION CO424322
 VERSION CO424322
 KEYWORDS GI:49670481
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@ds-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
 1..672
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"

/db_xref="taxon:10090"
 /clone="IMAGE:30665389"
 /tissue_type="whole eye"
 /dev_stage="newborn (1, 5, 15 days)"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH_BMAP_H10"
 /note="Organ: Eye; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AATAATTAAG. This library was created for the University
 of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene
 Discovery in the Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH)."

ORIGIN

Query Match 35.5%; Score 672; DB 7; Length 672;

Best Local Similarity 100.0%; Pred. No. 3.1e-164;
 Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 791 AGCAATCACTTAACGGGGCTTACCGTGAAGCCGCCAAGCCGTCGATGCTGCGCTGA 850
      |||||||
DB 1 AGCAATCACTTAACGGGGCTTACCGTGAAGCCGCCAAGCCGTCGATGCTGCGCTGA 60
      |||||||
QY 851 TCCGGTGGCTGTTAATGATGAACCCCAACCCGTCGGGCCCACTGAGAGATGTAGCCAGTC 910
      |||||||
DB 61 TCCGGTGGCTGTTAATGATGAACCCCAACCCGTCGGGCCCACTGAGAGATGTAGCCAGTC 120
      |||||||

```

FEATURES	source
LOCUS CK781308	
DEFINITION CK781308 655 bp mRNA linear EST 23-FEB-2004	
ACCESSION U1-M-G10-clp-g-23-0-U1.r1 NIH BMAP_g10 Mus musculus cDNA clone	
VERSION IMAGE:30619078 5', mRNA sequence.	
KEYWORDS CK781308.1 GI:42746986	
SOURCE EST.	
ORGANISM Mus musculus (house mouse)	
REFERENCE Mus musculus	
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE 1 (bases 1 to 655)	
JOURNAL NIH-MGC http://mgc.nci.nih.gov/ .	
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
CONTACT: Robert Strausberg, Ph.D.	
EMAIL: rsra@bbs-riemail.nih.gov	
Tissue Procurement: Dr. James Lin, University of Iowa	
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa	
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa	
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mouse1.html	
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)	
Seg primer: pYX-5.	
Location/Qualifiers	
1..655	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6"	

ORIGIN

sites of the pT73D-Pact vector. Library went through one round of normalization, and was constructed in the laboratory of M. Bento Soares (University of Iowa)."

Query Match 34.1%; Score 646; DB 4; Length 646;
Best Local Similarity 100.0%; Pred. No. 1,9e-157; Indels 0; Gaps 0;
Matches 646; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193 ATACCTCTTCTGCGCTGGCAAGACGACCTTAAAGCTTCCGAAAGGCAATTCAGAAAA 1252
DB 1 ATACCTCTTCTGCGCTGGCAAGACGACCTTAAAGCTTCCGAAAGGCAATTCAGAAAA 60
QY 1253 AGTCTCTTACTCTGCTGAGGGAGGTACAGAGGACCTCTCAGAACTCAGACCGGTGCTG 1312
DB 61 AGTCTCTTACTCTGCTGAGGGAGGTACAGAGGACCTCTCAGAACTCAGACCGGTGCTG 120
QY 1313 ATACTCCAGGGAGACCTGCTGCTGATCCGCTCCGAAAGGAGGCACTCTTAA 1372
DB 121 ATACTCCAGGGAGACCTGCTGCTGATCCGCTCCGAAAGGAGGCACTCTTAA 180
QY 1373 AGTCTCCAGACGCGTAACTGTTACTACTCTCTCCAGAGCCGAGGAGTCTGGGAGAC 1432
DB 181 AGTCTCCAGACGCGTAACTGTTACTACTCTCTCCAGAGCCGAGGAGTCTGGGAGAC 240
QY 1433 TCTTAAAGCGCAGTATGTGTTTGTGAGTGGGAGCCCGGTGAGCAGAACTCTCCAGAG 1492
DB 241 TCTTAAAGCGCAGTATGTGTTTGTGAGTGGGAGCCCGGTGAGCAGAACTCTCCAGAG 300
QY 1493 CTTCAGAGGCTCTCTCTCCAGCAGCAAGGCAATTCGAAACTCAATGAGGAGTCTCCGCA 1552
DB 301 CTTCAGAGGCTCTCTCTCCAGCAGCAAGGCAATTCGAAACTCAATGAGGAGTCTCCGCA 360
QY 1553 CAGCCTTAAAGAGGCACTACCCCTAGCACTTTGAGCTCCCTGAGCACTAGGCTCTCTCC 1612
DB 361 CAGCCTTAAAGAGGCACTACCCCTAGCACTTTGAGCTCCCTGAGCACTAGGCTCTCTCC 420
QY 1613 ATCTGAGAGCCCGGCGCAGCGCCCTCAGAGGCTGTGAGTGAAGACAGCACTCTGTCT 1672
DB 421 ATCTGAGAGCCCGGCGCAGCGCCCTCAGAGGCTGTGAGTGAAGACAGCACTCTGTCT 480
QY 1673 CCGAGTCTTGAACCAATTGGAATCTGCTGAAAGTCTCCGAAACCCCACTGAGGGAGCT 1732
DB 481 CCGAGTCTTGAACCAATTGGAATCTGCTGAAAGTCTCCGAAACCCCACTGAGGGAGCT 540
QY 1733 GTGTGTCTGTGAACAACCTGAGGGGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAGGAT 1792
DB 541 GTGTGTCTGTGAACAACCTGAGGGGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAGGAT 600
QY 1793 GGTGGCAGGAATCTTGGGGATGAGTGTCTTTCTGACAGACTG 1838
DB 601 GGTGGCAGGAATCTTGGGGATGAGTGTCTTTCTGACAGACTG 646

RESULT 14
BF789245 948 bp mRNA linear EST 12-JAN-2001
LOCUS 602105144F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223267
DEFINITION 5', mRNA sequence.
ACCESSION BF789245
VERSION BF789245.1 GI:12094281
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
plate: LHAM811 row: g column: 12
High quality sequence stop: 651.
Location/Qualifiers
1. 948
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4223267"
/lab_host="DH10B (T1 phage-resistant)"
/clone_11b="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORE6; Site:1; NCI;
Site:2; Salt: Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 34.1%; Score 645.8; DB 2; Length 948;
Best Local Similarity 96.3%; Pred. No. 2,4e-157; Indels 4; Gaps 2;
Matches 683; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

QY 867 GGTGAACCCCAACCGTCGCGGCACTGAGAGTGAAGCAGTATGAGGAGTCAACTG 926
DB 12 GTGAACCCCAACCGTCGCGGCACTGAGAGTGAAGCAGTATGAGGAGTCAACTG 71
QY 927 GGGTTACACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 986
DB 72 GGGTTACACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 131
QY 987 TGACTTGGCGGCGCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
DB 132 TGACTTGGCGGCGCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 191
QY 1047 GAATGAGCCCAAGGTGAGCTCTTCAAGACAGCAGTCCGAGAGTGAAGCACTGT 1106
DB 192 GAATGAGCCCAAGGTGAGCTCTTCAAGACAGCAGTCCGAGAGTGAAGCACTGT 251
QY 1107 ACCTGAGCTGAGACCGGCAATCTCTTAAAGAGTCCGAAAGAGATGAGTGA 1166
DB 252 ACCTGAGCTGAGACCGGCAATCTCTTAAAGAGTCCGAAAGAGATGAGTGA 311
QY 1167 AATCTGCAAGTGAACCGGCTGAGATCTTCTGCGCTGAGCAAGGAGCTTAA 1226
DB 312 AATCTGCAAGTGAACCGGCTGAGATCTTCTGCGCTGAGCAAGGAGCTTAA 371
QY 1227 GCTTCGAAAGGCAATCTCAAGAAAAAGTCTTACCTGCTGAGGAGGAGTACAGAGGA 1286
DB 372 GCTTCGAAAGGCAATCTCAAGAAAAAGTCTTACCTGCTGAGGAGGAGTACAGAGGA 431
QY 1287 CCCTCAGAACTCAGACCGGTGCTGATCTCAGGAGCAACCTGCTCCTGATCTCT 1346
DB 432 CCCTCAGAACTCAGACCGGTGCTGATCTCAGGAGCAACCTGCTCCTGATCTCT 491
QY 1347 GCTTCGAAAGGAGCAATCTTAAAGATCTGACAGCGTGAATCTGTTACTACTCTCT 1406
DB 492 GCTTCGAAAGGAGCAATCTTAAAGATCTGACAGCGTGAATCTGTTACTACTCTCT 551
QY 1407 TCCAGAGCCAGGAGTCTGGGAGCTCTTAAAGCAGCAGGAGTGTGTTGAGTGGAGGA 1466
DB 552 TCCAGAGCCAGGAGTCTGGGAGCTCTTAAAGCAGCAGGAGTGTGTTGAGTGGAGGA 611
QY 1467 CCCCGTGAAGCAGAGTCTTCAAGGCTTCAAGGCTCTCTTCAACCGCAAGGCAATCT 1526
DB 612 CCCCGTGAAGCAGAGTCTTCAAGGCTTCAAGGCTCTTCAACCGCAAGGCAATCT 669
QY 1527 CAAACTCAATGCGAAGTCTCCGCGCAGGCTTAAAGAGGAGTACCTCT 1575

Db 670 CAA-CTCATGCGAAGTCTCCGCGACAGCTTAAGCTACTAGCACT 716

RESULT 15
CA319312
LOCUS
DEFINITION
CA319312 795 bp mRNA linear EST 09-JUL-2003
IMAGE:6816569 5', mRNA sequence.
CA319312
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 795)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source
Seq primer: pyx-5.
Location/Qualifiers
1..795
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6816569"
/issue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1kb="NIH_BMAP_FW0"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 33.8%; Score 640; DB 6; Length 795;
Best Local Similarity 96.5%; Pred. No. 7.6e-156;
Matches 685; Conservative 0; Mismatches 0; Indels 25; Gaps 2;

QY 1 ATGAGTGGTGGCTTACTTCAGAGCGCCGAGCCAGGCTCCCTGGCTCCGCTGGCC 60
DB 81 ATGAGTGGTGGCTTACTTCAGAGCGCCGAGCCAGGCTCCCTGGCTCCGCTGGCC 140
QY 61 TCGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTGCCTTAACCTGTATGAAG 120
DB 141 TCGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTGCCTTAACCTGTATGAAG 200
QY 121 AAGCAGCGGTGAAGCGGACCATCAACAAACAACTGCGGACGCTAGAGTTCTTG 180

Db 201 AAGCAGCGGTGAAGCGGACCATCAACAAACAACTGCGGACGCTAGAGTTCTTG 260
QY 181 GAGACGCTGGGCAAGGCGACCTACGGAGGTGAAGAGCAGAGAGCTCGGAGCGT 240
DB 261 GAGACGCTGGGCAAGGCGACCTACGGAGGTGAAGAGCAGAGAGCTCGGAGCGT 320
QY 241 CTGGTGGCCATCAAGTGCATCAGAGAAAGACAAATCAAGATGACGAGTCTGTCAC 300
DB 321 CTGGTGGCCATCAAGTGCATCAGAGAAAGACAAATCAAGATGACGAGTCTGTCAC 380
QY 301 ATACGAGGGGAGATTGATCATGTCTTCACTCAACACCCCAATCATTTGCCATCAT 360
DB 381 ATACGAGGGGAGATTGATCATGTCTTCACTCAACACCCCAATCATTTGCCATCAT 440
QY 361 GAA-----GTGTTGAGATGACAGCAAGTGTGATTTGTC 396
DB 441 GAAGTGGCAGATCAAGCTTAGTACGTGTGTTTGAAGATGACAGAAATTTGATTTGTC 500
QY 397 ATGAGATGATCCAGCGGAGCGCATCTGTATGATTTACATCATGAGCGGCGCATGAGT 456
DB 501 ATGAGATGATCCAGCGGAGCGCATCTGTATGATTTACATCATGAGCGGCGCATGAGT 560
QY 457 GACGGGAGCGGAGCGCATTTCTTCCGACAGATGTGTCTGCTGCACTACTGCAACAG 516
DB 561 GAGGGGAGCGGAGCGCATTTCTTCCGACAGATGTGTCTGCTGCACTACTGCAACAG 620
QY 517 AAGGGATCGTTACCGGAGATCAAGCTGGAACATCCTCTAGATGCAATGGAAC 576
DB 621 AAGGGATCGTTACCGGAGATCAAGCTGGAACATCCTCTAGATGCAATGGAAC 680
QY 577 ATCAGATTGCTGATCTTGGCTCTCCACCTGTACCAAGGCAAGTTCTTCAGACG 636
DB 681 ATCAGATTGCTGATCTTGGCTCTCCACCTGTACCAAGGCAAGTTCTTCAGACG 740
QY 637 TTCTGTGGAGCGCTCTTACGCTCGGCTGAGATGATCAACGGGAAGCC 686
DB 741 TTCTGTGGAG-CCTCTTACGCTCGGCTGAGATGATCAACGGGAAGCC 789

Search completed: May 16, 2005, 03:15:29
Job time : 6076.42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:12:18 ; Search time 162.033 Seconds
(without alignments)
1994.176 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293
Sequence: 1 MESVALLQRPSPQAPSALA.....DCQEVTAAYRQALGICSLKS 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trernbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	100.0	631	2 Q9DBV0	Q9DBV0 m mus muscu
2	3285	99.8	631	2 Q8CIC0	Q8CIC0 mus musculus
3	3279	99.6	639	2 Q80ZM4	Q80ZM4 mus musculus
4	3272	99.4	639	2 Q8BZM4	Q8BZM4 mus musculus
5	3087.5	93.8	630	2 Q6GHE5	Q6GHE5 rat mus muscu
6	2824.5	85.8	638	2 Q9H093	Q9H093 homo sapien
7	1949.5	59.2	438	2 Q6BEO4	Q6BEO4 homo sapien
8	1699.5	51.6	661	1 ARK5 HUMAN	Q60285 homo sapien
9	1695.5	51.5	658	2 Q641K5	Q641K5 mus musculus
10	1515.5	46.0	575	2 Q616D6	Q616D6 mus musculus
11	1008.5	30.6	461	2 Q8CGE1	Q8CGE1 mus musculus
12	888	27.0	1551	2 Q65ZM4	Q65ZM4 caenorhabdi
13	888	27.0	1592	2 Q7KSS0	Q65ZM3 caenorhabdi
14	881.5	26.8	1180	2 Q7KSS0	Q7KSS0 drosophila
15	881.5	26.8	1180	2 Q7YU19	Q7YU19 drosophila
16	881.5	26.8	1427	2 Q9VH05	Q9VH05 drosophila
17	853	25.9	383	2 Q7QBH4	Q7QBH4 anopheles g
18	748	22.7	752	1 MRK4 HUMAN	Q96134 homo sapien
19	745	22.6	752	1 Q8CIP4	Q8CIP4 mus musculus
20	743	22.6	752	2 Q8NG37	Q8NG37 homo sapien
21	716.5	21.8	792	2 Q6TNT7	Q6TNT7 xenopus lae
22	701	21.3	729	2 Q9JKE4	Q9JKE4 mus musculus
23	701	21.3	744	2 Q9JKE5	Q9JKE5 mus musculus
24	700.5	21.3	797	2 Q8VHF0	Q8VHF0 rat mus muscu
25	696.5	21.2	780	2 Q804T1	Q804T1 xenopus lae
26	695.5	21.1	993	2 Q7KSK4	Q7KSK4 drosophila
27	693.5	21.1	795	2 Q8VHJ5	Q8VHJ5 mus musculus
28	693	21.0	778	2 Q96HB3	Q96HB3 homo sapien
29	692	21.0	725	2 Q804T2	Q804T2 xenopus lae
30	691.5	21.0	793	2 Q08678	Q08678 rat mus muscu
31	691.5	21.0	795	2 Q9P0L2	Q9P0L2 homo sapien

32	691	21.0	712	2 Q7QBH3	Q7QBH3 anopheles g
33	690.5	21.0	755	2 Q7KZ17	Q7KZ17 homo sapien
34	690.5	21.0	1098	2 Q7KXK7	Q7KXK7 drosophila
35	688.5	20.9	888	2 Q8BR95	Q8BR95 mus musculus
36	688	20.9	745	2 Q15524	Q15524 homo sapien
37	686.5	20.8	774	1 MRK2 MOUSE	Q05512 mus musculus
38	685.5	20.8	705	2 Q9VGF9	Q9VGF9 drosophila
39	684.5	20.8	771	2 Q69Z17	Q69Z17 mus musculus
40	684	20.8	832	2 Q963B6	Q963B6 drosophila
41	680.5	20.7	719	2 Q6B8A18	Q6B8A18 homo sapien
42	680	20.6	776	2 Q7ZYL7	Q7ZYL7 xenopus lae
43	680	20.6	785	2 Q8GCV3	Q8GCV3 xenopus lae
44	678.5	20.6	691	2 Q96RC0	Q96RC0 homo sapien
45	678.5	20.6	834	2 Q8WVX1	Q8WVX1 haemochus

ALIGNMENTS

RESULT 1
Q9DBV0
ID Q9DBV0 PRELIMINARY; PRT; 631 AA.
AC Q9DBV0:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,
clone:1200013822 product:weakly similar to PROBABLE SERINE/THREONINE-
PROTEIN KINASE KIAA0537 (EC 2.7.1.-) (Mus musculus adult male cecum
cDNA, RIKEN full-length enriched library, clone:910215K18
DE CDNA, RIKEN full-length enriched library, clone:910215K18
DE product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
DE KIAA0537).
GN Name=1200013822R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning."
Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto S., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.",
RL Genome Res. 10:1757-1771(2000).

RP [6]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takamashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.

RP [7]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saesaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takamashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AK036737; BAB23518.1; -
DR HSSP; P31751; 1GZK.
DR MGD; MGI:1921387; 1200013B22Rik.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002230; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00117; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 631 AA; 69806 MW; D51C042DECC174B CRC64;

Query Match 100.0%; Score 3293; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 6e-181;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLRPSQAPASALASASAPLADGLTKSPKLMKKQAVRHHKHLRRHYEFL 60
DB 1 MESVALLRPSQAPASALASASAPLADGLTKSPKLMKKQAVRHHKHLRRHYEFL 60
QY 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAIIH 120
DB 61 ETLGKTYGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPHIIAIIH 120

QY 121 EYFENSSKTVIMEYASRGDLYDYTSERPRLSEPPARHFRROIYSALHYCHONGIVHRDL 180
DB 121 EYFENSSKTVIMEYASRGDLYDYTSERPRLSEPPARHFRROIYSALHYCHONGIVHRDL 180
QY 181 KLENILLDANGIKIADFGLSNLYKKGKPLQTFCCSPLYASBEIVNGKRYVPEVDSMSL 240
DB 181 KLENILLDANGIKIADFGLSNLYKKGKPLQTFCCSPLYASBEIVNGKRYVPEVDSMSL 240
QY 241 GVLVLYLVHGTPFPGQDHKTLVKQISNGAYNEPPKPSDACGLIRMLMVNPTRRATLED 300
DB 241 GVLVLYLVHGTPFPGQDHKTLVKQISNGAYNEPPKPSDACGLIRMLMVNPTRRATLED 300
QY 301 VASHMVMVMGTYTTGCEQGEALRECGHPSGDERRASMAWMLRRSSRPILENGAKVCSFFQ 360
DB 301 VASHMVMVMGTYTTGCEQGEALRECGHPSGDERRASMAWMLRRSSRPILENGAKVCSFFQ 360
QY 361 HYPGGGSTYPLGERHSLKSKRKENDMAONLOGDPAEDTSSRPKSSSLKPLKGLKKSS 420
DB 361 HYPGGGSTYPLGERHSLKSKRKENDMAONLOGDPAEDTSSRPKSSSLKPLKGLKKSS 420
QY 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLIPKKGILKKSRORESGYSSPEPSESGELD 480
DB 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLIPKKGILKKSRORESGYSSPEPSESGELD 480
QY 481 ASDVFSVSGDPVQKSPQASGLLHKRGLIKNGKRSRRLBECTTSTPSLQGLASSHPA 540
DB 481 ASDVFSVSGDPVQKSPQASGLLHKRGLIKNGKRSRRLBECTTSTPSLQGLASSHPA 540
QY 541 ARPSRPSGAVSDSTLSSESFPQDLPERLPETPLRGCVSVNLTGLEOPSPSEGLKRWQ 600
DB 541 ARPSRPSGAVSDSTLSSESFPQDLPERLPETPLRGCVSVNLTGLEOPSPSEGLKRWQ 600
QY 601 ESLGDSCEFLTDQEVTAAYRQALGICSKLS 631
DB 601 ESLGDSCEFLTDQEVTAAYRQALGICSKLS 631

RESULT 2
Q8CIC0 PRELIMINARY; PRT; 631 AA.
AC Q8CIC0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
GN 1200013B22Rik protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary tumor;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strusberg R.L., Reitgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.J., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC033302; AAH3302.1; -.
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 631 AA; 69745 MW; C6C904224D9174F CRC64;

Query Match 99.6%; Score 3285; DB 2; Length 631;
 Best Local Similarity 99.8%; Pred. No. 1.7e-180;
 Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESVALLORPSQAPSAALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 DB 1 MESVALLORPSQAPSAALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 QY 61 ETLGKTYGKVKARSSGRLVAIKSRKDKIDEODLHRRREIEMSSLNPHITAIH 120
 DB 61 ETLGKTYGKVKARSSGRLVAIKSRKDKIDEODLHRRREIEMSSLNPHITAIH 120
 QY 121 EVFENSKIYIWEVYASRGDLYISRRPRLSRDANHPROQVSAHYHONGIYARDL 180
 DB 121 EVFENSKIYIWEVYASRGDLYISRRPRLSRDANHPROQVSAHYHONGIYARDL 180
 QY 121 EVFENSKIYIWEVYASRGDLYISRRPRLSRDANHPROQVSAHYHONGIYARDL 180
 DB 121 EVFENSKIYIWEVYASRGDLYISRRPRLSRDANHPROQVSAHYHONGIYARDL 180
 QY 181 KLENIILLDANGNIKIDFGLSNLYHKGFQTFCGSPLVSPETVNGKPYVGEVDSMSL 240
 DB 181 KLENIILLDANGNIKIDFGLSNLYHKGFQTFCGSPLVSPETVNGKPYVGEVDSMSL 240
 QY 181 KLENIILLDANGNIKIDFGLSNLYHKGFQTFCGSPLVSPETVNGKPYVGEVDSMSL 240
 DB 181 KLENIILLDANGNIKIDFGLSNLYHKGFQTFCGSPLVSPETVNGKPYVGEVDSMSL 240
 QY 241 GVLLYILVHGTMPPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 DB 241 GVLLYILVHGTMPPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 QY 241 GVLLYILVHGTMPPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 DB 241 GVLLYILVHGTMPPDQDHTLVKQISNGAYREBPKEPSACGLIRWLLMVPTRATLED 300
 QY 301 VASHMVMWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRPLLENGAKVCSFFKQ 360
 DB 301 VASHMVMWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRPLLENGAKVCSFFKQ 360
 QY 301 VASHMVMWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRPLLENGAKVCSFFKQ 360
 DB 301 VASHMVMWNGYITTVGVEQELAREGHHSGPFGASAMDMRRSRPLLENGAKVCSFFKQ 360
 QY 361 HVGGGSTVPGELRQHSLLKSRKENDMAQMLQGDPAEDTSRRPKSKSLKPKGILLKKSS 420
 DB 361 HVGGGSTVPGELRQHSLLKSRKENDMAQMLQGDPAEDTSRRPKSKSLKPKGILLKKSS 420
 QY 361 HVGGGSTVPGELRQHSLLKSRKENDMAQMLQGDPAEDTSRRPKSKSLKPKGILLKKSS 420
 DB 361 HVGGGSTVPGELRQHSLLKSRKENDMAQMLQGDPAEDTSRRPKSKSLKPKGILLKKSS 420
 QY 421 TSSGEVQEDPQELRPVDPTRGQVPVAVSLPRKGIILKKSQRBSGYSSPEPSSEGLLD 480
 DB 421 TSSGEVQEDPQELRPVDPTRGQVPVAVSLPRKGIILKKSQRBSGYSSPEPSSEGLLD 480
 QY 421 TSSGEVQEDPQELRPVDPTRGQVPVAVSLPRKGIILKKSQRBSGYSSPEPSSEGLLD 480
 DB 421 TSSGEVQEDPQELRPVDPTRGQVPVAVSLPRKGIILKKSQRBSGYSSPEPSSEGLLD 480
 QY 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALLEGTPSTFGSLDOLASSHPA 540
 DB 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALLEGTPSTFGSLDOLASSHPA 540
 QY 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALLEGTPSTFGSLDOLASSHPA 540
 DB 481 ASVFPVSGDVEKQSPQASGLLHRRKILLKNGFSTALLEGTPSTFGSLDOLASSHPA 540
 QY 541 ARSPRSGAVSEDSIISSEFPDQLDPERLPETPLRCVSVNDLRGLQEPSEGLKRWQ 600
 DB 541 ARSPRSGAVSEDSIISSEFPDQLDPERLPETPLRCVSVNDLRGLQEPSEGLKRWQ 600
 QY 541 ARSPRSGAVSEDSIISSEFPDQLDPERLPETPLRCVSVNDLRGLQEPSEGLKRWQ 600
 DB 541 ARSPRSGAVSEDSIISSEFPDQLDPERLPETPLRCVSVNDLRGLQEPSEGLKRWQ 600
 QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
 DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

DB 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 3
 0802M3
 ID 0802M3 PRELIMINARY; PRT; 639 AA.
 AC 0802M3;
 DT 01-JUN-2003 (TEMBLrel. 24, Created)
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE 1200013B22Rik protein.
 GN Name=1200013B22Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klanner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heien F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huylk S.W.,
 RA Vailion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
 RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC046833; AAH46833.1; -.
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser_Thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 639 AA; 70675 MW; FB9C40228F53872C CRC64;

Query Match 99.6%; Score 3279; DB 2; Length 639;
 Best Local Similarity 98.7%; Pred. No. 3.9e-180;
 Matches 631; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MESVALLORPSQAPSAALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60
 DB 1 MESVALLORPSQAPSAALASASARPLADGLIKSPKLMKQAVKRRHHKNTLRHYEFL 60


```

QY 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120
DB 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120
QY 121 E-----VFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQ 172
DB 121 EVGRSLVTVFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQ 180
QY 173 NGIVHBDLKLLENLILANGNIKADFGLSNLVHKGFQLQFCGSPVYASPEIYNGKPYVG 232
DB 181 NGIVHBDLKLLENLILANGNIKADFGLSNLVHKGFQLQFCGSPVYASPEIYNGKPYVG 240
QY 233 PEVDNSLSGLVLLVILVHGTMPFDGQDKTLVKQISNGAVAREPPKPSDACGLIRWLLMVP 292
DB 241 PEVDNSLSGLVLLVILVHGTMPFDGQDKTLVKQISNGAVAREPPKPSDACGLIRWLLMVP 300
QY 293 TRRATLEVDVASHMVMWVGTTTGVEQBALREGHPSGDPRASMDMLRRSSRPILLENGA 352
DB 301 TRRATLEVDVASHMVMWVGTTTGVEQBALREGHPSGDPRASMDMLRRSSRPILLENGA 360
QY 353 KVCSPFKOHVPGGSGTVPGLEROHSLKKSREKENDMAQNLQGDPAEDTSSRPGKSLKLPK 412
DB 361 KVCSPFKOHVPGGSGTVPGLEROHSLKKSREKENDMAQNLQGDPAEDTSSRPGKSLKLPK 420
QY 413 GILKKKSTSSGVEQDPOLRFPVPTPGQPVPAVSLPRKGIKKSRQRESGYSSPSP 472
DB 421 GILKKKSTSSGVEQDPOLRFPVPTPGQPVPAVSLPRKGIKKSRQRESGYSSPSP 480
QY 473 SESGELLDSADVFSGDVPEQKSPQASGLILHRKGIKLNGKSRRLBETTPSTGSLD 532
DB 481 SESGELLDSADVFSGDVPEQKSPQASGLILHRKGIKLNGKSRRLBETTPSTGSLD 540
QY 533 QLAASHPARSPRPSGAVSDESLSSSPDQDLPERLPETPLRGCVSVNLLGLGLOPPS 592
DB 541 QLAASHPARSPRPSGAVSDESLSSSPDQDLPERLPETPLRGCVSVNLLGLGLOPPS 600
QY 593 EGLKRWMOESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 601 EGLKRWMOESLGDSCFSLTDCQEVTAAYRQALGICSKLS 639

```

RESULT 5
Q66HES PRELIMINARY; PRT; 630 AA.

AC 066HES;
AT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DS Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Alteschul S.F., Zeeberg B., Buettner K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusika K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavatoni T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McManus P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Phelan J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttenfield Y.S.,
RA Krzyzanski M.I., Skolala U., Smalun D.E., Scherch A., Schein J.E.,

```

RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Director MGC Project;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL, BC081899; AA081899.1; -  
DR InterPro: IPRO01009; Kinase-like.  
DR InterPro: IPRO00719; Prot_kinase.  
DR InterPro: IPR002290; Ser_thr_kinase.  
DR InterPro: IPR008271; Ser_thr_kinase.  
DR InterPro: IPR001245; Tyr_kinase.  
DR Pfam: PF00069; Kinase; 1.  
DR ProDom: PD000001; Prot_kinase; 1.  
DR SMART; SM00220; S_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.  
DR PROSITE; PS0011; PROTEIN KINASE_DOM; 1.  
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.  
KW Hypothetical protein.  
SQ  
SEQUENCE 630 AA; 69952 MW; 1626427ABCD5F66E CRC64;  
  
Query Match 93.8%; Score 3087.5; DB 2; Length 630;  
Best Local Similarity 94.3%; Pred. No. 3.8e-169;  
Matches 595; Conservative 13; Mismatches 22; Indels 1; Gaps 1;  
  
QY 1 MESVALLORPSQAPSASALASBARPLADGLIKSPKLMKQAVKHHKHNIRHYEFL 60  
DB 1 MESVALLORPSQAPSASALASBARPLADGLIKSPKLMKQAVKHHKHNIRHYEFL 60  
QY 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120  
DB 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120  
QY 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120  
DB 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIIMSLNPHIIAIH 120  
QY 121 EVFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIVHRLD 180  
DB 121 EVFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIVHRLD 180  
QY 121 EVFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIVHRLD 180  
DB 121 EVFENSSKTIIVMEYASRGDLVDYISERPRLSERARHPFRQIVSALHYCHQIVHRLD 180  
QY 181 KLENILLDANGNIKADFGLSNLVHKGFQLQFCGSPVYASPEIYNGKPYGVDVSWL 240  
DB 181 KLENILLDANGNIKADFGLSNLVHKGFQLQFCGSPVYASPEIYNGKPYGVDVSWL 240  
QY 241 GVLLYTLVHGTMPFDGQDKTLVKQISNGAVAREPPKPSDACGLIRWLLMVPTRATLED 300  
DB 241 GVLLYTLVHGTMPFDGQDKTLVKQISNGAVAREPPKPSDACGLIRWLLMVPTRATLED 300  
QY 301 VASHMVMWVGTTTGVEQBALREGHPSGDPRASMDMLRRSSRPILLENGAVCSFFKQ 360  
DB 301 VASHMVMWVGTTTGVEQBALREGHPSGDPRASMDMLRRSSRPILLENGAVCSFFKQ 360  
QY 361 HYPGGSGTVPGLEROHSLKKSREKENDMAQNLQGDPAEDTSSRPGKSLKLPKILKKSS 420  
DB 361 HYPGGSGTVPGLEROHSLKKSREKENDMAQNLQGDPAEDTSSRPGKSLKLPKILKKSS 420  
QY 421 TSSGVEQDPOLRFPVPTPGQPVPAVSLPRKGIKKSRQRESGYSSPSP 480  
DB 421 TSSGVEQDPOLRFPVPTPGQPVPAVSLPRKGIKKSRQRESGYSSPSP 480  
QY 481 ASDVPFSPGVPVROKSPQASGLILHRKGIKLNGKSRRLBETTPSTGSLD 540  
DB 481 ASDVPFSPGVPVROKSPQASGLILHRKGIKLNGKSRRLBETTPSTGSLD 540  
QY 541 ARSPRPSGAVSDESLSSSPDQDLPERLPETPLRGCVSVNLLGLGLOPPSGGLKRW 600  
DB 541 ARSPRPSGAVSDESLSSSPDQDLPERLPETPLRGCVSVNLLGLGLOPPSGGLKRW 600  
QY 601 ESTGDSGSLTDCQEVTAAYRQALGICSKLS 631  
DB 601 ESTGDSGSLTDCQEVTAAYRQALGICSKLS 630

```


RC TISSUE=Fetal kidney;
 RG The German cDNA Consortium;
 RA Othenwelder B., Obermaier B., Deutschenbaur S., Schaap A.,
 RA Mewes H.W., Weill B., Amd C., Osanger A., Fobio G., Han M., Wiemann S.,
 RA Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CH749209; CH18066.1;
 DR InterPro: IPR001009; Kinase_like.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Prodom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Hypothetical protein.
 KW NON TER
 FT
 SQ SEQUENCE 438 AA; 47528 MW; 6C5E0DB89D4D5C4 CRC64;
 Query Match 59.2%; Score 1949.5; DB 2; Length 438;
 Best Local Similarity 84.6%; Pred. No. 5.3e-104;
 Matches 373; Conservative 17; Mismatches 44; Indels 7; Gaps 2;

195 IADFGSNLYHKGKFLQTFGSPLYASPRIVNGKPVYGEVDSWSLGLVYLIVHGTMP 254
 1 IADFGSNLYHKGKFLQTFGSPLYASPRIVNGKPVYGEVDSWSLGLVYLIVHGTMP 60
 255 DGDHNTLVKQISNGAYRPPKPSDACGLIRWLMVNPTRATLEVDVASHWVVMGYTG 314
 61 DGDHNTLVKQISNGAYRPPKPSDACGLIRWLMVNPTRATLEVDVASHWVVMGYATR 120
 315 VEEQELALRGHSGDPGASAMADWLRRSSRPLLENGAYCSPFKOHVPGGSTVGLER 374
 121 VEEQELALRGHSGDPGASAMADWLRRSSRPLLENGAYCSPFKOHVPGGSTVGLER 180
 375 QHSLLKSRKRNMAOMLQGDPAEDTSRPGKSLKLPKGLIKKSSSTSGEVDEDEQER 434
 181 QHSLLKSRKRNMAOMLQGDPAEDTSRPGKSLKLPKGLIKKSSSTSGEVDEDEQER 240
 435 PVPDTFGQVPAVSLPRKGLIKKSRQSSGYSSPEPSBSGGLDASIVFVSGDVEOK 494
 241 PVPDTFGQVPAVSLPRKGLIKKSRQSSGYSSPEPSBSGGLDASIVFVSGDVEOK 297
 495 SPQASGLILHRKGLIKLNGKFSRTALEGTPSTFGSLDOLASSHPARPSPGAVSEDS 554
 298 SPQASGLILHRKGLIKLNGKFSRTALEGTPSTFGSLDOLASSHPARPSPGAVSEDS 357
 555 ILSSESFDOLDLPERLPETPLRGCVVDNLRGLEQSPSSG---LKRMOESLGDSCFL 610
 358 ILSSESFDOLDLPERLPETPLRGCVVDNLRGLEQSPSSGCLRMWDPLGDSGCSFL 417
 611 TDQCEVTAAYRQALGICSKLS 631
 418 TDQCEVTAAYRQALGICSKLT 438

RESULT 8
 ARKS HUMAN
 ID ARKS HUMAN STANDARD; PRT; 661 AA.
 AC 060285;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE AMPK-related protein kinase 5 (EC 2.7.1.37).
 GN Name=ARKS; Synonyms=KIAA0537;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagae T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. IX.

RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 RN [2]
 RP CHARACTERIZATION, AND MUTAGENESIS OF SER-600.
 RX MEDLINE=22393479; PubMed=12409306; DOI=10.1074/jbc.M206025200;
 RA Suzuki A., Kusaka G.-I., Kishimoto A., Lu Y., Ogura T., Lavie M.F.,
 RA Esnault H.;
 RT "Identification of a novel protein kinase mediating Akt survival
 RT signaling to the ATM protein."
 RL J. Biol. Chem. 278:48-53(2003).
 CC -1- FUNCTION: Involved in tolerance to glucose starvation.
 CC Phosphorylates ATM.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- ENZYME REGULATION: Activated by PKB/AKT1 during glucose
 CC starvation.
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in heart and brain,
 CC and at lower levels in skeletal muscle, kidney, ovary, placenta,
 CC lung and liver.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB011109; BAA25463.1; -.
 DR HSSP: P31751; IGZK.
 DR MIM: 608130; -.
 DR InterPro: IPR011009; Kinase_like.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR Pfam: PF00069; Pkinase; 1.
 DR Prodom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Phosphorylation; Polymorphism;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 55 306
 FT NP BIND 61 69
 FT ACT SITE 178 178
 FT BINDING 84 84
 FT MOD RES 600 600
 FT VARIANT 543 543
 FT FT
 FT MUTAGEN 600 600
 SQ SEQUENCE 661 AA; 74304 MW; 806F37D52CA4718F CRC64;
 Query Match 51.6%; Score 1699.5; DB 1; Length 661;
 Best Local Similarity 55.5%; Pred. No. 2.1e-89;
 Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

14 PSASALASASAPRLDGLIKSPKPLKQAVRHHKHLRRYELTLTGKTYKVK 73
 13 PDLGIGAPSPBRAVAGATPALEP-RKPGVKKHHKHLKRYELTLTGKTYKVK 71
 74 ARBS-SGRVVAIKSIKDKIKDEODLHTRREIEMSLNHPHIIAHEVFENSXIVY 132
 72 ATERPSGRVVAIKSIKDKIKDEODLHTRREIEMSLNHPHIIAHEVFENSXIVY 131
 133 MEYASRGDLYDYISERRPRLSEDAHFFRQIVSALHYCHONGIVHBDLKLENIILDANGN 192
 132 MEYASRGDLYDYISERRPRLSEDAHFFRQIVSALHYCHONGIVHBDLKLENIILDANGN 191
 193 IKIADFGSNLYHKGKFLQTFGSPLYASPRIVNGKPVYGEVDSWSLGLVYLIVHGTMP 252
 192 IKIADFGSNLYHKGKFLQTFGSPLYASPRIVNGKPVYGEVDSWSLGLVYLIVHGTMP 251
 253 PFDGQDHNTLVKQISNGAYRPPKPSDACGLIRWLMVNPTRATLEVDVASHWVVMGYAT 312

```
DB 252 PFGFPHKQULIRISSGEYREPTQPSARGLIRMLMVNDRAITIEDIANHVMVWGKY 311
QY 313 TGVGBQDALREGBPGSGDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGGSIVPGL 372
DB 312 SSVCDCCDALHDSSEPL-----LARIIDWHNRSTGLQADTEAKMGKLAK---PTTSEVW--L 362
QY 373 ERGHSLLKSRKENDMAQNLQGDPAEDTSSRPGKSLKPKGILKKKS-----STSGEVQ 427
DB 363 ERGRSLKSKKENDPQSGQDAVPS-----PSKLSSRRPGILKKKSNSHRHSHTGFIR 418
QY 428 -----EDPQEL-----RPVDTPGQPPAVS--LLPRKGLKKRSORESGY 466
DB 419 GVVGPALPTFKMEQDLCTRGVLLPSSPEAFVPGKLSFKOSATMPKKGILKTKQORESGY 478
QY 467 YSSPESSEGEILDADVFVSGDPVEQKSP-----QASGLLHRKGLIKLNGKFSRTAL 520
DB 479 YSSPESSESEILLDSNDVM--GSSIPSPSPDPARVTSLSLSCRKGLIKHSSKYSAGTM 526
QY 521 EGTTPS-----TGSLLQ-----LASSHAARSPSPGAVSDESLSSSEFPQDLPER 569
DB 537 DPALVSPKPTLESISEPGVAPAGLSRSY-----SRSSVTSDDSVLSSDSPDLLOEN 591
QY 570 LP-ETPLRGCSVDN-----LRGLBOPSEGLKRWMOESLGDSCFS-LTDCQEVTAAYR 621
DB 592 RPAPGRIRSCVSAENFLQIDFEGLOKRRPQYLKXIRKLADSSFSLLTMDMDVTQVTK 651
QY 622 QALGICSKLS 631
DB 652 QALGICSKLN 661

RESULT 9
Q61KS PRELIMINARY; PRT; 658 AA.
AC 0641K5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE RIKEN CDNA B231014P22.
GN Name=B231014P22R1K;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strusberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Pfrange C.,
Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahney J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyniaki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
```

```
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC082328, AA092328.1, -.
SQ SEQUENCE 658 AA; 73661 MW; E7176f3b75b899f CRC64;
Query Match 51.5%; Score 1695.5; DB 2; Length 658;
Best local similarity 56.0%; Pred. No. 3.5e-89;
Matches 371; Conservative 77; Mismatches 136; Indels 79; Gaps 20;
QY 18 ALASESARPLADGLIKSPRLMKQAQVKKHHKHNLRHYEFLETIGKTYGKVKARAS 77
DB 26 AVAGATAADV-----EPR---KPHGVKHHKHNKHYELQETIGKTYGKVKATAT 76
QY 78 -GRVVAISIKDKTKDEODLLHTRREIEMSSLNHPHITAIHFENFSSKVIYMEYA 136
DB 77 FSGRVVAISIKDKTKDELDMVHTRREIEMSSLNHPHISIYEFENKDKVIIMEYA 136
QY 137 SRGDLVDYISERPRISERDARHFFROIYALHYCHONGVHVDLKIENITLDANGIKIA 196
DB 137 SKGELYDYISERRRISERTRHFFROIYSAVHCHONGVHVDLKIENITLDNCNIIKIA 196
QY 197 DFGLSNLYHKGRFLQTFCCGSPLYASPEIVNGKRYVGEVDSVSLGVLTYLVHGTMPFDG 256
DB 197 DFGLSNLYGKDKFLQTFCCGSPLYASPEIVNGKRYVGEVDSVSLGVLTYLVHGTMPFDG 256
QY 257 QDHKTIVKQISNGATREPPKESDAGLIRWLLMVPTRATLEDVASHHVMVWGTYTGC 316
DB 257 FDHKNLIRQISSGEYREPTQPSDARGLIRWLLMVPTRATLEDIANHVMVWGYSYVC 316
QY 317 EOBALREGBPGSGDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGGSIVPGERQH 376
DB 317 DCDALPDSSEPL-----LARIIDWHNRSTGLQADTEAKMGKLAK---PQASEVW--LERQR 367
QY 377 SLKSKRKENDMAQNLQGDPAEDTSSRPGKSLKPKGILKKKS-----STSGEVQ--- 427
DB 368 SLKSKSKENDPQSGQDSVPS-----PSKLSSRRPGILKKKSNSHRHSHTGFIEGIVS 423
QY 428 ---EDPQEL-----PVPTPGQVPA--VSL-----LPRKGLIKRSORESGYSSP 470
DB 424 PALPSPFKMEQDLCTAIFLPSSPEADWSGKLSIQSATMPKKGILKTKQORESGYSSP 483
QY 471 EPSSEGEILDADVFVSG-----DPVEQKSPQASGLLHRKGLIKLNGKFSRTALGTPSS 526
DB 484 EHSSESEILLDSNDVVISGGLSPPDPARGTSHLSCKRKGLIKKSSRSRD---GGTDA 540
QY 527 TFGSLDQLASSHPARP-----SRPSGAVSDESLSSSEFPQDLPERLP-ETPLR 576
DB 541 LTRPEMPTLESISPQVPSDGISRSYSRPSIISDVSLSDFSFLLEQENRAPRQRIR 600
QY 577 GCVSVDNLKGL-----QPSSEGLKRWMOESLGDSCFS-LTDCQEVTAAYRQALGTS 628
DB 601 SCVSAENFLQIDFETPHNRPAPQYLK-----LADSSFSLLTMDMDVTQVYKALGICS 655
QY 629 KLS 631
DB 656 KLN 658

RESULT 10
Q61D6 PRELIMINARY; PRT; 575 AA.
AC 061D6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MKIAA0537 protein (Fragment).
GN Name=mkIAA0537;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hirooka S.,
```

RA Saga Y., Hagase T., Ohara O., Koga H.;
 RT "Prediction of the Coding Sequences of Mouse Homologues of KIA Gene:
 RT III. The Complete Nucleotide Sequences of 500 Mouse KIA-Homologous
 RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
 RT Randomly Sampled from Size-fractionated Libraries.",
 RL DNA Res. 10:167-180(2003).
 RN (2)
 RN SEQUENCE FROM N.A.
 RA Okazaki N., Kikuno R., Nagase T., Ohara O., Koga H.;
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL: AB182364; EMB23895.1; -;
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: GO:0016740; P:transferase activity; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR011009; Kinase_like.
 DR InterPro: IPR00719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00219; TKC; 1.
 DR SMART: SM00219; TKC; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON_TER
 SQ SEQUENCE 575 AA; 64949 MW; 3CA0796FB4199D2B CRC64;
 Query Match 46.0%; Score 1515.5; DB 2; Length 575;
 Best Local Similarity 55.7%; Pred. No. 6.4e-79;
 Matches 332; Conservative 70; Mismatches 125; Indels 69; Gaps 17;
 QY 84 IKSIRKDKIKDEODLHIREIREIMSLNPHIATHEVEENSKIVIMEYASRGDLYD 143
 DB 1 IKSIRKDKIKDEODLHIREIREIMSLNPHIATHEVEENSKIVIMEYASRGDLYD 60
 QY 144 YISERPRLSERDARHFRQIVSAMATYCHQNGIVYHRLKLENIILLDANGIKINDPGLSNT 203
 DB 61 YISERPRLSERDARHFRQIVSAMATYCHQNGIVYHRLKLENIILLDANGIKINDPGLSNT 120
 QY 204 YHKGKFLQTFCCGSPLYASPEIVNGKPYGVGVDSWISGLVLLYLVGTMFPDGDHKTLY 263
 DB 121 YQDKKFLQTFCCGSPLYASPEIVNGKPYGVGVDSWISGLVLLYLVGTMFPDGDHKTLY 180
 QY 264 KOISNGAYREPPKPSDAGCGLIRWLMLVNPTRATLEADVASHMVMNMGYTTGVGEORALNE 323
 DB 181 ROISSEVEYREPTQPSDARGILRMMLVNPTRATLEADVASHMVMNMGYTTGVGEORALNE 240
 QY 324 GGHPSGDPRASMAWMLRRSSRPDLLENKAVCSFFKQHVPGGISTVPGLERQHSKKSRK 383
 DB 241 SESPL-----LARIIDWHHRSTGLQAEAEAKMKGLAK---FGASEV--LEKQSLKSKK 291
 QY 384 ENDMAQNLGDRPREDSSRRPGKSLKLPGLIKKKS-----STSSSEVQ-----EDPG 431
 DB 292 ENDPFGSGDVSPE-----PSKLSRRPKGILKKRNSSEHRSHSTGTIGIVPALPSP 347
 QY 432 ELR-----PYVDTFGQVPVA-VSL-----LPRKGIKKSRQRESGYVSPESSEGR 477
 DB 348 KHEQDLCRTAIPSPSPADMSGKLSLQSAATMPKGIKKTORRESGYVSPESSEGR 407
 QY 478 LLDASVFWVG---DPVEQKSPQASGLILHAKGILKLGAKSFRTALEGTPS-TPGSLD 532
 DB 408 LLDNSNVVVISGLSSPPPPARCTSHSLSCRRKGIKLGHSRYSD---GGTDPALTPPEM 464
 QY 533 QLASHPAARP-----SRPGVASEDSIILSESPDOLDLPERLP-ETPLAGCVSYND 583
 DB 465 TLESLSPPGVSPGICISRSYISRPSSITSDSVLSSDSFDLDELQENPARQRTISCVSAEN 524
 QY 584 LRGLE-----OPPSBGLKRWMOBSLGDSCFS-LTDCQEVTAAYRQALGICSKLS 631

DB 525 FLOLQDFETPHNRPPOYLKR-----LADSPSFLITMDMDVQVYKKALETCSKLN 575
 RESULT 11
 Q8CGE1
 ID Q8CGE1 PRELIMINARY; PRT; 461 AA.
 AC Q8CGE1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE B230104P22Rik protein (Fragment).
 GN Name=B230104P22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;
 RN (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=salivary gland;
 RC PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaeetz T.E.,
 RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kerteman M., Madan A.C., Rodichenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.B.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RN SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=salivary gland;
 RA Strausberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC040467; AAH40467.1; -;
 DR MGI: MGI:1925226; B230104P22Rik.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004672; F:protein kinase activity; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR011009; Kinase_like.
 DR InterPro: IPR00719; Prot_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 FT NON_TER
 SQ SEQUENCE 461 AA; 51274 MW; 0522EC4ECF3713BB CRC64;
 Query Match 30.6%; Score 1008.5; DB 2; Length 461;
 Best Local Similarity 49.0%; Pred. No. 6.3e-50;
 Matches 236; Conservative 59; Mismatches 118; Indels 69; Gaps 17;
 QY 198 FGLSNLYHKGKFLQTFCCGSPLYASPEIVNGKPYGVGVDSWISGLVLLYLVGTMFPDQ 257
 DB 1 FGLSNLYHKGKFLQTFCCGSPLYASPEIVNGKPYGVGVDSWISGLVLLYLVGTMFPDQ 60
 QY 258 DHKTLVQKISNGAYREPPKPSDAGCGLIRWLMLVNPTRATLEADVASHMVMNMGYTTGVGE 317
 DB 61 DHKTLVQKISNGAYREPPKPSDAGCGLIRWLMLVNPTRATLEADVASHMVMNMGYTTGVGE 120
 QY 318 QEALREGHPSGDPRASMAWMLRRSSRPDLLENKAVCSFFKQHVPGGISTVPGLERQHS 377

Db 121 CDALPSESPL-----LARIIDWHNRSTGLQABABAKKGLAK---PGASEV--LERQRS 171
Qy 378 LKSRKRENDMAOWLQGDPAEDTSSRPCKSLKPKGILKKKS-----STSGEVQ----- 427
Db 172 LKSKSKENFPQSGQSVSPS-----PSKSSKRPKGILKKRSSEHRSHSTGTGIVSP 227
Qy 428 --EDPOELR-----PVDPDTPGQPVPA-VSL-----LPRKGIILKKSRQREGGYSSPE 471
Db 228 ALPSPFKRMQDLCRTAIPLPSPSPADMSGSLKQSGATMKKGIILKKTKQRESGYTSSPE 287
Qy 472 PEEGGLLDASDVFSVSG---DVEQKSPQASGLLHKRGIILKNGKFSRTALEGTTPS- 526
Db 288 RESESLDLSNDVVISGLSSPPDPARGTSHSLGCRKGIILKHSRYSQ---GGTDPL 344
Qy 527 TFGSLDQLASSHPAAR-----SRPGAVSEDSILSSESPDQLDPERLP-ETPLRG 577
Db 345 TRPEMTLESLSPGVPVDSIGSRSPSPSSILSDSVLSSDSPDLLETQENRPARORIS 404
Qy 578 CVSDVPLRGLE-----OPSEGLKRMWQESLSDSCFS-LTDCOEVTAAVRGALGICSK 629
Db 405 CVSAEFLQLQDPEETHNRPRPOLKR-----LADSSFLLTMDVDTQVYKPALEICSK 459
Qy 630 LS 631
Db 460 LN 461

RESULT 12
Q65ZH4 PRELIMINARY; PRT; 1551 AA.
ID 065ZH4;
AC 065ZH4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3;
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCB1_Taxid=6239;
RN 11
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology". The C. elegans Sequencing Consortium.";
RT Science 283:2012-2018(1998).
RN 12
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Le T.T.;
RT "The sequence of C. elegans cosmid B0496.";
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN 13
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN 14
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RT Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58749; AUT05597.1; -
DR InterPro; IPR011009; Kinase, like.
DR InterPro; IPR002280; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.

DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SF; 1.
KW Hypothetical protein.
SQ SEQUENCE 1551 AA; 176395 MW; 1F3BE62632A44C95 CRC64;
Query Match 27.0%; Score 888; DB 2; Length 1551;
Best Local Similarity 37.0%; Pred. No. 2,6e-42;
Matches 230; Conservative 73; Mismatches 214; Indels 104; Gaps 15;
Qy 54 RHRVEFLRTGKGTGKVKARESS-GRVVAIKSRKDKIKDQODLHTRRETISSIN 112
Db 72 KHRFEITKQSGSTGKSLAYDHKFDREVAALIKKSAIESKADLVRIREIRISALN 131
Qy 113 HPHIATHEVFENSSKIVMEYASRGDIYDISSERPLSERDARFFRQIVSALYCHQ 172
Db 132 HENIIDIYVEFNKKIILVMEYSSGGLYDYVSRGSLPEABEARIFRQIVSALYCHK 191
Qy 173 NGIVHDLKLENIILDPANGNIKIADPGLSNLYHKKFLTFPGSPPLIYASPEYNGKPYVG 232
Db 192 HRVAHRDLKLENIILDPQNNNAKIADFGLSNYFADKNLITTFGSPPLIYASPEYINGPYNG 251
Qy 233 PEVDSWSLGLVLYILVHGTMPDGDHKTLYKQISNGAYREPPKPSDAGLIRMLNVP 292
Db 252 PEVDCWSLGLIYLTVYGSMPDGRDPRNRMVQIKRGAIFEPETSTASMLIRMLNVP 311
Qy 293 TRRATLEDVASHMVMVNGYTTGV-----GEOBALREG 324
Db 312 ERRATIFDIASHMVLNLEENMPVIOELPENQIIDHTPLTERBETMIVQDLADQDVFMF 371
Qy 325 GHPSPDFGASAMDLR-----SRPPLIENAKYCSFKQIVPGGGSTV 369
Db 372 GHLSSETRKIKDFIRKKEAEFPNDNSPVKPKARKTDELGKISKEQPEEKSAKSL 431
Qy 370 PGLERQHS-----LKSRK-ENDMAOWLQGDPAEDTSSRPCKSLKPKGILKKK 418
Db 432 RGVKEEKEKPVVDNDPLERLQIENRIGQKKKQKAKTARAVETVKL-----EVK 485
Qy 419 SSTSGEVOEDPOELRPVDPDTPGQPVAVSLPRKGIILKKSRQREGGYSSPESEGL 478
Db 486 KEKSPQGPEDPKTKAR-----GTSKPADSRAP-SFVPIKQDPER---SEERPRTRH 534
Qy 479 LDASDVFSVSGDVEQKSPQASGLLHKRGIILKLN-GKFSRTALEGTTSTFGSLDQLAS 536
Db 535 LTPASAVRIETDLSLNMNMNV--LEQMEKGPVMIITARIKAPLVDTRPMVKELLESITA 592
Qy 537 SHPARPSRPGAVSEDSILSSESPDQ---DLPERLPETPLRGCVSDNLRGLQEPS 592
Db 593 AQPPEVQKQTSKVVEQQTFSRONTLTRKKGEDPLEEIPV-----SPS 637
Qy 593 EGLK-RWV-----QESLG 604
Db 638 RKMKRPMHVSVCMKNSVG 658
RESULT 13
Q65ZH3 PRELIMINARY; PRT; 1592 AA.
ID 065ZH3;
AC 065ZH3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3;
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCB1_Taxid=6239;
RN 11
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RX

RG WormBase Consortium;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2016(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br1etol N2;
 RT Murray J., Le T.T.;
 RT "The sequence of *C. elegans* coemid B0496.";
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br1etol N2;
 RA Waterston R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br1etol N2;
 RG WormBase Consortium;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U58749; AAU05598.1; -.
 DR Interpro: IPR011009; Kinase_like.
 DR Interpro: IPR000719; Prot_kinase.
 DR Interpro: IPR002290; Ser_thr_kinase.
 DR Interpro: IPR008271; Ser_thr_pkin_AS.
 DR Interpro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Kinase; 1.
 DR Prodom: PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 1592 AA; E28D6FF934C3F76 CRC64;
 Query Match 27.0%; Score 888; DB 2; Length 1592;
 Best Local Similarity 37.0%; Pred. No. 2,7e-42;
 Matches 220; Conservative 73; Mismatches 214; Indels 104; Gaps 15;

Db 535 LTRASAVRITDLSNMLNMOV--LEONEKGPVNIITARIKAMPDYTRPVMKELLSITA 552
 QY 537 SHPARPSPGCVSDESIILSSSEFPOL-----DLPERLPETPLRGCVSYNDNLGLRPPS 592
 Db 593 AGPEPVQKOTSKVVEQOTSRONTLTRKKKEDPLEEIEFVP-----SPS 637
 QY 593 EGLK-RMW-----QESLG 604
 Db 638 RKMKERPMHSVEVMKNGESVG 658
 RESULT 14
 ID Q7KS50 PRELIMINARY; PRT; 1180 AA.
 AC Q7KS50;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE CG11870-PD.
 GN ORFNames=CG11870;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxId=7227;
 RX [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.U., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R.K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Flook A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodok A., Gong F., Gorelli J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modyarty C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheel F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodagat, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhu Q., Zhu X., Smith H.O.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RX Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitskas R., Taber P.E., Man K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.",
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svitskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celisner S.E.,
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomic perspective.",
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk K.S.,
 RA Smith C.D., Tupy J.L., Whitfield E.D., Bayraktaroglu L., Bergman S.E.,
 RA Belencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.",
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG Flybase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AB003686; AAS6535.1; -.
 DR HSSP; Q63450; 1A06.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase-like.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KM KM
 SQ SEQUENCE 1180 AA; 127743 MW; A88A012664046EC6 CRC64;

Db 120 QIMSVHNNIIHYEFENRREKMLVMEFAAGELDYDLRSRKVLTEEARIRFQVAT 179
 Qy 166 ALHYCHQNGIYVRDLKLENTLIDANGNIKIADFGSLNHYKGFLOTGCGSLYASPEIV 225
 Db 180 AVYCHCHKIKICRDLKLENTLIDANGNIKIADFGSLNHYKGFLOTGCGSLYASPEIV 239
 Qy 226 NGKPYVGPVDSMSGLVLYLTHGTMPPDGDHLYVQISNGVAREPCKSDCGCLR 285
 Db 240 EGTPIVQGPVDSMSGLVLYLTHGTMPPDGDHLYVQISNGVAREPCKSDCGCLR 299
 Qy 286 WLLMNPTRATLEBDVASHMWN-----WGVTGVGEQALR-----E 323
 Db 300 DWLTVCPRRABATLEQCSHMWNENDNNSCLADBLANQTVRDVLLSLTPATTAAQ 359
 Qy 324 GGHPSGDFGRASMDWLRSSR-----PLENGAVYCSFKQHPGGG---- 367
 Db 360 LVVPSAEGAAAKAANERVPSSHVSIGIRDMGPPTTEARRLLDM-----VAAAGEAALM 415
 Qy 368 -----TVPLEGQHSLSKSKKENDMAQNIQGPADTSRPGKSL-----KLPK 412
 Db 416 PSPRTITPAQSPVOTKRLQPTVSTENAGTTAK-KKEKPANSSFVISKDAPLLEADP 474
 Qy 413 GILKKSSSTSGE-VOEDPOLRPVPD-----TPGQPVPA-V 447
 Db 475 TIEBQATLMEAFITANIPREV-PVPSYSQKMDAVGDCILMGATTPAPAPAPPTV 533
 Qy 448 SLIPKRG-----ILKK----- 458
 Db 534 AQPTRGKLDVAVETPEEKDAATKVIKKFVNKHKTADLVNAINESAKAAPVAVAPPF 593
 Qy 459 -----SRQRESGYSSPEPESGELLDAIDVPVSGPVPKQSQ-ASGLLHKKGLKNG 513
 Db 594 VRKCSLQDSTLKNKNAERRKSRILETAEKFPPPPPVAAAPPEKPKKLSIPGVSVGSPFK 653
 Qy 514 KFSR-----TALEGTTPSTFGSLDQLASSHPAR-----PSRPSGAVSSEDSILSSSS 560
 Db 654 EFKKATNPAPAEGETPGLTRAQEQVAAAAAAGAEATLSTPSPVVAQSLGSDSKKS 713
 Qy 561 FQDLDPPE 568
 Db 714 VASISLIDE 721
 RESULT 15
 Q7YU19 PRELIMINARY; PRT; 1180 AA.
 AC Q7YU19;
 DT 01-OCT-2003 (TRENBERG 25, Created)
 DT 01-OCT-2003 (TRENBERG 25, Last sequence update)
 DT 01-MAR-2004 (TRENBERG 26, Last annotation update)
 DE LP05937P.
 GN ORFNames=CG11870;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Broksstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Fattán D., Frise E.,
 RA George R., Gonzalez M., Guarín H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.,
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BT010033; AAC22502.1; -.
 DR HSSP; P31751; 1MRV.
 DR Flybase; FBgn0037804; CG11870.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

Job time : 167.033 secs

DR GO:0016740; P:transferase activity; IEA.
DR GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR Pfam: PF00069; PKinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 1180 AA; 127684 MW; 2CD790A4F3D8C50C CRC64;

Query Match 26.8%; Score 881.5; DB 2; Length 1180;

Best Local Similarity 32.7%; Pred. No. 4,3e-42;

Matches 238; Conservative 101; Mismatches 216; Indels 173; Gaps 21;

QY 6 LQRP-SQAPSASALASASARPLADL-----IKSPKPLMKQA-----VKR 46
DB 2 VSKPDGTAPNGAGAAAP--TGLDATGNSLHPSGIPDQIDNMSGTANTGNVKK 59
QY 47 HHHKNTLRHYEFLTYGKTYGKKA-RESSGLVAIKSRKDKIKDQDLHIREI 105
DB 60 NNRKKLRORFDIIKKLGQGYGVQVLGINKEGQVAIKTKICKIEADLVIRREV 119
QY 106 EIMSSLNHPHIIAHEVFENSSKIVMEVYASGDLVDYISERPLSERDARHFFROIVS 165
DB 120 QIMSSVHHNPIIHYVFENREKMLVMEFPAAGGELYDYISERKVLTEEARIRFQVAT 179
QY 166 ALHYCHONGIVHRDLKLENIILLDANGNIKIADFGSLNLYHKGKFLQTPCGSPLYASPEIV 225
DB 180 AVYCHKHKICHRDLKLENIILLDEKNAKIADFGLSNVFDDQRLSTPCGSPLYASPEIV 239
QY 226 NKPVYGPPEVDSMSLGVLLYILVHGTPPPDGDQDKTLVQISNGAYRPPPKPSDACGLIR 285
DB 240 EGPYQGPPEVDCMSLGLVLLYTLVYSGMPDGSNFKRLVQISQGDYRPPPKPSRASTLIR 299
QY 286 WILLNPTRRATLIEDVASHMVN-----WGTTGVGEQEAR-----E 323
DB 300 DMLTVCRRRASIEOICSHMVVENDNVSCLDLAEDLANOTPVRLVLLSLTPATTADQ 359
QY 324 GHPSGDFGRASMDWLRSSR-----PLENGAKVCSFQKHVPGGS----- 367
DB 360 LVVPSAEGAALKAAANERVPRSHVSGSTRDMGPPTTEARRILDM---VAAGGAALM 415
QY 368 ----TPGLERQHSLSKSRKENDMAQNTQGDPAEDTSRPGKSL-----KLPK 412
DB 416 PSPTRTITPAQSPVQTKRKLOPTVSTENNAAGTTAK-KKEKPANSSFVISKDGAPLTEAPP 474
QY 413 GILKKSSSTSSGE-VEEDQELRPVPD-----TRGQVPVA-V 447
DB 475 TIIEPQATIMEAETIANIPEEV-FVPSYSOKMQAVGDCIDLMGTATTAPAPAPPTPV 533
QY 448 SLPRKG-----ILKK----- 458
DB 534 AROPTRGKLDAYVETPEEKDATKVIKKFVNHKKTADLVANINESAKAAAPVSAVAPPP 593
QY 459 ----SRQRESGYVSSPEPSGELDASDVVFGSDPEVQKSPQ-ASGLLHRKGIILKNG 513
DB 594 VRKCSLQDSTLNKFNAAERKRSRLTETAEKFGPPVAAAPPEKPKKLSIPGVSGSPKK 653
QY 514 KFSR-----TALGTTPTSTGSLDQILASSHPAAK-----PSRPGAVSEDSIILSSS 560
DB 654 EPEKATNPAAAGPTPGELRAQGVAAAAAAQAEBETLSTPPSPVVAQSLGSDSKNS 713
QY 561 FDQLDLP 568
DB 714 VASISIDE 721

Search completed: May 11, 2005, 14:24:00

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:13:08 ; Search time 38.5452 Seconds
(without alignments)
1575.104 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293

Sequence: 1 MESVALIQRPSQAPSASALA.....DCQEVTAAYRQALGICSKLS 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	888	27.0	1558	2 T29253	hypothetical prote
2	694.5	21.1	713	2 S27966	probable serine/ch
3	689	20.9	745	2 G01025	serine/threonine p
4	686.5	20.8	774	2 T14609	probable serine/ch
5	669	20.3	1192	2 T18611	-probable serine/ch
6	667	20.3	1246	2 G89287	protein H39E23.1 (
7	664	20.2	1398	2 T13741	hypothetical prote
8	649.5	19.7	512	2 T52633	serine/threonine-s
9	648.5	19.7	633	1 A26030	serine/threonine-s
10	648	19.7	798	2 JC7500	qlk protein - chic
11	641	19.5	504	2 T10449	probable serine/ch
12	639.5	19.4	512	2 JCI446	serine/threonine-s
13	631	19.2	511	1 A56009	serine/threonine-s
14	630.5	19.1	552	1 S51025	[hydroxymethyl]glut
15	627.5	18.9	602	2 S72513	FOG2 protein - yea
16	623.5	18.9	552	1 A53621	[hydroxymethyl]glut
17	622	18.9	562	2 T29858	hypothetical prote
18	621	18.9	472	2 B90100	SNF-related kinase
19	620	18.8	504	2 T07415	probable serine/ch
20	612	18.6	481	2 T49072	protein kinase
21	610	18.5	473	1 S53941	serine/threonine-s
22	599	18.2	512	2 T07788	probable serine/ch
23	596.5	18.1	887	2 T20941	hypothetical prote
24	594.5	18.1	576	2 T41587	probable carbon ca
25	587.5	17.8	622	1 S44859	serine/threonine-s
26	583.5	17.7	891	2 A38903	protein kinase 1 -
27	581.5	17.7	891	2 T40503	protein kinase kin
28	579	17.6	510	2 T04145	serine/threonine p
29	570	17.3	502	1 A41361	serine/threonine-s

30	569.5	17.3	401	2 B90120	SNF1-related prote
31	564	17.1	513	1 S60303	serine/threonine-s
32	563	17.1	513	1 S60304	serine/threonine-s
33	559	17.0	1147	2 S64930	serine/threonine-s
34	549.5	16.7	441	2 S84667	probable protein k
35	549	16.7	1064	2 S52687	serine/threonine-s
36	547	16.6	1142	2 S53959	G1N4 protein - yea
37	545	16.6	651	2 S52244	p69Eg3 protein - A
38	543.5	16.5	480	2 A86427	probable serine/th
39	543	16.5	800	2 S29344	protein kinase KIN
40	539.5	16.4	440	2 T14735	probable serine/th
41	537.5	16.3	440	2 T14736	probable serine/th
42	534.5	16.2	435	2 E84707	probable protein k
43	529.5	16.1	726	2 T33998	hypothetical prote
44	527	16.0	746	2 S62365	SNF1-related prote
45	524	15.9	502	2 T02306	probable protein k

ALIGNMENTS

RESULT 1	
T29253	hypothetical protein B0496.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans	
C:Date: 15-Oct-1999	#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29253	
R:Murray, J.; Le, T.T.	
Submitted to the EMBL Data Library, May 1996	
A:Description: The sequence of C. elegans cosmid B0496.	
A:Reference number: Z20596	
A:Accession: T29253	
A>Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: DNA	
A:Residues: 1-1558 <MDR>	
A:Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3	
A:Experimental source: strain Bristol N2; clone B0496	
C:Genetic8:	
A:Gene: CESP:B0496.3	
A:Map position: 4	
A:introns: 55/1; 100/3; 141/1; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2;	
Query Match	27.0%; Score 888; DB 2; Length 1558;
Best local similarity	37.0%; Pred. No. 2.1e-28;
Matches	230; Conservative 73; Mismatches 214; Indels 104; Gaps 15;
QY	54 RRRYFLETLGKGYGKVKARSS-GRIVAISIRKDKIKDQDILHRRREIINSGLN 112
DB	72 KRRFEITTKLGGTYGKSLAYDHKFDREAVAKLKKSAIESKADIVRIRREIRIMSALN 131
QY	113 HHIIAIHVEFENSRIYVMEYASRGDIYDYSRPRLESDAHHFFQIYSALHYCHQ 172
DB	132 HNNIIQIYEVFENKCKIILVMEYSSGGELYDVASRCGSIPLEAEARIFQIYSALVYCHK 191
QY	173 NGIVRDLKLENIILDANGNIKIADFGSLNLYHKKGFLOTFCGSPLYASPEIYNGKPYVG 232
DB	199 HVAHRDLKLENIILDONNNAKIADFGSLNRYADKMLLTFFCGSPLYASPEIYNGPYVG 251
QY	233 PEVDWSLGLVLLYLIVHGTMPDGDHKTLYQJISNGAYREPPKPSDAGLIRMLLVNP 292
DB	252 PEVDWSLGLIILYLYVSGMPDGRDFNRMVQIKRGAFEEBETSTASMLIRNMLRVNP 311
QY	293 TTRATLEDVASHWVNMWGTITGV-----GEQKLRG 324
DB	312 ERRATIFDLASHWMLNLEENMPVIOELPENQIIDHTPLTEREETMIVODLADQDFWMP 371
QY	325 GHPSGDFGRASWADLRR-----SSRPLENKAQYCSFFKQHVPGGQSTV 369
DB	372 GHLSEBTRKIKIDPFIIRKEAEFPNDNSPVKPKARKDELNGKISKQPEEMKAKSKSL 431
QY	370 PGLERQHS-----LKSRK-ENDMAQNLQGDPAEDTSRPSKSLKPKGLKKK 418
DB	432 RGVKEKEKPKVVDNDPLERLROJENRLGQNKQKKAATKTSARVETVTLK-----EVK 485

Qy	419	SSTSGGEVQEDQPEQLRPVPDTPGQVPVPAVSLPRKGI LKKSQRKRESGSYSSPEPSESSEL	478
Db	486	KEKSEQPEQDEPKTAR-----GTSKRPADSRAP-SFVVKQRPET-----SEPERPRTPPH	534
Qy	479	LDADVPFVSGPVEQKSPQASGLILHRKGI LKLN--GKFSRTALEGTTPTFGSLDQLAS	536
Db	535	LTAASVRIETSSLNTLMNQV--LEQMEKGPVNLTINARIKHPLYDTPTMYKELIESITIA	592
Qy	537	SHPAAPSRPSPGAVSEDSILSSSEFDQ-----DLPERLPETPLRGCVSVNDIRGLDEQPS	592
Db	593	AQPEPVQKQTSKVVEQQTFPSRQNTLTTRKKKEDPLEEELPEV-----SPS	637
Qy	593	EGLK-RWV-----QESLG	604
Db	638	RKMKERPMHSAVEVCMKNEISVG	658

RESULT 2

N:Alternative names: protein p78
 C:Species: Homo sapiens (man)
 C:Idate: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
 C:Accession: S27966
 R:Mareshwari, K.K.; Som, S.; Parsa, I.
 submitted to the EMBL Data Library, January 1992
 A:Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced
 A:Reference numbers: S27966
 A:Accession: S27966
 A:Molecule type: mRNA
 A:Residues: 1-713 <MAH>
 A:Cross-references: UNIPROT:P27448; EMBL:M80359; NID:g189511; PIDN:AA55991.1; PID:g1895991
 C:Superfamily: protein kinase homology
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 E:154-307/Domain: protein kinase homology <KIN>
 E:62-70/Region: protein kinase ATP-binding motif

[illegible]

OY	500	GILHRK-----GILKNGKFRALAGTTPSTGSLDQASH---PAAP	543
		:::::	:
Dd	484	GGMTRRNTVYCSERTTDDRHVSYLQ-NGKENSTIPORT-----VASTSHISSAAPT	534
OY	544	SR---PSGAIVEDS	554
		:	:
Dd	535	DRIRFPGTAARST	548

RESULT 3

serine/threonine protein kinase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C:Accession: G01025
R:Navarro, E.
submitted to the EMBL Data Library, April 1996
A:Reference number: H00564
A:Accession: G01025
A:Status: preliminary; translated from GB/EMBL/DBJ
A:molecule type: mRNA
A:Residues: 1-745 <NAV>
A:Cross-references: UNIPROT:O15524; UNIPROT:Q96HB3; EMBL:X57630; NID:91310674
A:Superfamily: protein kinase homology
P18-271/Domain: protein kinase homology <KIN>

Query Match	20.9%	Score 689	DB 2	Length 745
Best Local Similarly	33.2%	Pred. No. 8.4e-21		
Matches 189	Conservative 92	Mismatches 176	Indels 112	Gaps 20

```

OY 57EELFETJGKGYGVYKAKRE--SSGGLVAKISRKOKIDOEODLIHREIETMSLNHH 115
Db 20YRLTKTJGKGNPAKVKARHILTLGVEAVKIIDKTL--NSSLOKLFREVRIMKVLNNH 78
OY 116IIAHEFVENSCKIYIVMEYASRGDLXYDISERPLTSERDANHFPROIYSALHYCHONI 175
Db 79IVKLFVEIETEKTLYLWMEYASGGEVFPYILAHGMMKEKAPAKRQIVSNVQYCHQKI 138
OY 176VHRDKLENIILLDANGNIKIADFGLSNLVYHKGFLQTFQCSBPYASPEIUVNGKPYVPEV 235
Db 139VHRDLKLENIILLDDAMNIKIDFGFSNEFTGNKLDTCGSPPYAPELFGQKXVDGPEV 198
OY 236DSWELGVLTYLIVNGTFMPDQDOKHTLYVKQISNGAYRREPCKPSDC--GLIRULMLNPR 294
Db 199DWSLGVLTLYLVSGSLFPDQONKELERVLRGCRYLPYWSJTCENILKKFLILNPK 258
OY 295RATEDVASHVMVMGYTTVGEOALREGGHPSGDFPRASMAWDLRRSSRPLENGAVY 354
Db 259RGTEFQIMKRMWNVGH-----EDDELKPYVEPLPDYDP-----RTELMWSMG--- 303
OY 355CSFPRQHVPGGGSITVPLGERQHSLKKSRENDMA-----ONLQGD-----PAED 398
Db 304--YRREI-----ODSLVGQRYNEVMATYLLGYKSSLEBGGTITLKRPSPAD 349
OY 399--TSSRPGKSSILKPKGI-----LKKSSTSJSGE--VOEDPOE 432
Db 350LTNSSAOPFSHKYQSVASANKORRFPDQAPAIPTSNSYSKKTOSNAENKRPBEDRES 409
OY 433LRPVPDT--PGOPVPAVSLLPKGLIKKSRQRESGYSSPSPBSGGLDADSVFVSGD 489
Db 410GRKASSTAKVPAJPLP-----GLEKK-----TTPPSTNSVLSTIN----- 447
OY 490PVEOKSPQASGLILHRKGIKTL--NGKFSRTALEGTPSPFGSILQOLASSHPAARPSK 546
Db 448-RSNNSP-----LLERASLGQASIQNGDOSTLT-MGSAASASTASAGVSAANPRQHQSM 500
OY 547SGAV--SEDSILSSEFPDQDLPERLPE 572
Db 501SGSVHPNKASGLPPTESNCEVPRSTAPQ 529

```

RESULT 4

148609
Probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #ext_change 16-Aug-2004
C:Accession: 148609; S31333
R:Inglis, J.D.; Lee, M.; Hill, R.E.
Mamm. Genome 4, 401-403, 1993
A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
A:Reference number: 148609; PMID:93364122; PMID:8358177
A:Accession: 148609
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-774 <ING1>
A:Cross-references: UNIPROT:Q05512; EMBL:X70764; NID:G57919; PIDN:CAAS0040.1; PID:G57920
R:Inglis, J.D.; Lee, M.; Hill, R.E.
submitted to the EMBL Data Library, January 1993
A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome
A:Reference number: S31333
A:Accession: S31333
A:Molecule type: mRNA
A:Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTSGTCYAGAHQATRT', 731-774 <ING2>
A:Cross-references: EMBL:X70764
C:Genetics:
A:Gene: emk
C:Superfamily: protein kinase homology
C:Keywords: ATP, phosphotransferase, serine/threonine-specific protein kinase
F:51-304/Domain: protein kinase homology <KIN>
F:59-67/Region: protein kinase ATP-binding motif

Query Match 20.8%; Score 686.5; DB 2; Length 774;
Best Local Similarity 32.6%; Pred. No. 1.1e-20;
Matches 199; Conservative 90; Mismatches 218; Indels 103; Gaps 20;

6 LLDPSQAPASALASASARPLADGLIKSPKPLMKQAVKRRHKNLHNR-----YEEL 60
11 LNERTEOPTLGLHDS-----KPSKSNMGRKRSATSDQPHIGNYRL 56
61 ETLGKGYGVKKVKAARE-SSGRVLAIKSIRKDKIKDEODLLHIREIEMSLNHPHII 119
57 KTIQGNFAKVLAAHILILGKEVAVKIIDKTQL-NSSLOKLFREVRIMKVLNHPHIVL 115
120 HEVFNSSKIIVIMEYASRGDLVDYISERPLSRDARHPFQIYSAHYCHONGIVHD 179
116 FEVITEKTLVIMEYASRGVEVDYLVAGRMKEKRAKRFQIVLVQYCHQKFIYHD 175
180 LKLENIILDANGNIKIADFGSLNLYHKKGFLOTFCSSPLVASPEIVNGRPYVGPEDS 239
176 LKAENILLADADNMIKIADFGFNSLFFGKLDITFCGSPYAAPBELFQGGKIDGPEVD 235
240 LGLVLYLVHGTMPPDGODHKTLLVKQISNGAYREPPKPSDAC-GLIRMLMVPTRATL 298
236 LQVILTYLVSGSLPFDGQNLKELREYVLAGKTRIPRYMTDCENILKFLILNPSRGTL 295
239 EDVASHMVMVWGTYTGVGEQALREGHPSGDFGRASMDWLRSSRPULLENGAKVCS 358
236 EQIMDRMWN-----VGHED-----DELKPYEPLLTG 325
339 KQHVGGGSGTVFGL---ERQHSIKSRKKNDAQNL-----OGDPAEDTSSRPKSSLT 410
336 RDRVVG---VNGLTTEEIODSLVGRVYEVAVATYLLGYKSEBPGDITTLKPRSSADL 381
411 PGILIKKSSSTSSGVEQDPOELRPVDPTRGQVPVAVSLPRKGLIKSRQESGYYSP 470
382 TNSASPSBHKVQSVSANPKQRR---SSDAVPMI---PTSNSYKSKTQSNNAENKRP 434
471 EPESEGLLDASDVFGDP-VEQK---SPQASGL-----LHHRKGLIKL- 511
435 ESETERKASSTAKVPASPLPGDLRKKXTTAPSTNSVLTSTNRSRSLPDLDRASLQASI 494
512 -NGKTSKTLLEGTTBTFGSLDQLASHP-----AARPRPGAVSEDSILSSSEF 561
495 QNGKDSLT-MPGSRASTASASAVSARPRQHKQMSASVHPRKASGLPPTES--NCEVP 551

QY 562 DQDLPERP 571
DB 552 RPSRAPORVP 561

RESULT 5
T18611
Probable serine/threonine-specific protein kinase (EC 2.7.1.-), long splice form - Caen
N:contains: probable serine/threonine kinase, short splice form
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C:Accession: T18611; T18610; T23144; T23143
R:McMurray, A.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z18997
A:Accession: T18611
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W113>
A:Cross-references: UNIPROT:Q9TW45; EMBL:Z81027; PIDN:CAAS4179.1; GSPDB:GN00023; CESP:H3
A:Experimental source: clone AH10
A:Accession: T18610
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W112>
A:Cross-references: EMBL:Z81027; PIDN:CAAS4178.1; GSPDB:GN00023; CESP:H39E23.1b
A:Experimental source: clone AH10
R:McMurray, A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z19696
A:Accession: T23144
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1192 <W113>
A:Cross-references: EMBL:Z96102; PIDN:CAAS4263.1; GSPDB:GN00023; CESP:H39E23.1a
A:Experimental source: clone H39E23
A:Accession: T23143
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-487, 536-1192 <W114>
A:Cross-references: EMBL:Z96102; PIDN:CAAS4262.1; GSPDB:GN00023; CESP:H39E23.1b
A:Experimental source: clone H39E23
C:Genetics:
A:Gene: CESP:H39E23.1a; CESP:H39E23.1b
A:Map position: 5
A:Insertions: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 98/
C:Keywords: alternative splicing; ATP, phosphotransferase, serine/threonine-specific pr
F:1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #
F:1-487, 536-1192/Product: probable serine/threonine-specific protein kinase, short splic

Query Match 20.3%; Score 669; DB 2; Length 1192;
Best Local Similarity 32.4%; Pred. No. 8.2e-20;
Matches 187; Conservative 97; Mismatches 218; Indels 76; Gaps 16;

11 SQAPSASALASASARPLADG-----LKSPLMKQAVKRRHKNLHNR 58
113 SGAPPAAS-CGSSAATVSSGSRHPTSGSSSHARSTGCGMSRSARANDDVHNGK 171
59 FLETLGKGYGVKKVKAARE-SSGRVLAIKSIRKDKIKDEODLLHIREIEMSLNHPHII 117
172 LKTIQGNFAKVLAAHILILGKEVAVKIIDKTAL-NPSLOKLFREVRIMKVLNHPHIVL 230
118 ALHEVFNSSKIIVIMEYASRGDLVDYISERPLSRDARHPFQIYSAHYCHONGIYH 177
231 KLYQVMEETQTLVLEVASGGEVVDYLVAGRMKEKRAKRFQIVSAVQYLSKNITL 290
178 RDLKLENIILDANGNIKIADFGSLNLYHKKGFLOTFCSSPLVASPEIVNGRPYVGPED 237
291 RDLKLENIILDQNMIMIKIADFGFNSLFFGKLDITFCGSPYAAPBELFQGGKIDGPEVD 350
238 WSLGVLYLVHGTMPPDGODHKTLLVKQISNGAYREPPKPSDAC-GLIRMLMVPTRAT 296
351 WSLGVLYLVSGSLPFDGQNLKELREYVLAGKTRIPRYMTDCENILKFLILNPSRGTL 410

```
QY 297 TLEDVASHWVWVWGYTTTGVGEQALREGGHPSD-----FGASMDW 339
D 411 SLDNIMKORWMMVWGY-----EDDELKPFLEPPDQIDECRIEKLQIFOLGFKKALILS 465
QY 340 LRRSSRPILLENGAKVCSFFKOHVPGGGSTVPGLEROHSLKKSRKENDMAONLQGDPAEDT 399
D 466 VKE-----KREDIATYLLG-----ERKSDMDAS--BITMAOSLSHSSINV 507
QY 400 SSRPKSLKLPKGLKKK---SSTSSEVQEDPQELRPVPTPGQPPVAVSLPRKGL 456
D 508 SSSLQOH-----PAGVITREHVTSSASGSSASPSRYSRSATATGASITAGSLASANA 563
QY 457 KKSRRQRESYSSPSESSEGLLDASDVVSGDPVEQKSPQASGLLIH-----RKGIKL 511
D 564 QHQOQSSAPSSSGSSSSSRSSQNDAA-ATAAGGVWMSGTRHGQVQMRAPTSRQATISL 622
QY 512 NGKFSRTALEGTT-----PSTFGSLDQGLASSHPAARPS 544
D 623 LQPPSYKSSNTTQIAQIPPLFNRSSTATSS--AAQPS 658
```

RESULT 6

```
G89287
protein H39E23.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: G89287
R:Anonymous, The C. elegans Sequencing Consortium.
S:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A15000; MIMD:99069613; PMID:9651916
A>Note: see webites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A>Note: published erittra appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 <STO>
A:Cross-references: GB:chr_V; PIDN:CAB09532.1; PID:g3878100; GSPDB:GN00023; CESP:H39E23.
C:Gene: H39E23.1
A:Map position: 5
```

```
Query Match 20.3%; Score 667; DB 2; Length 1246;
Best Local Similarity 31.9%; Pred. No. 1e-19;
Matches 182; Conservative 93; Mismatches 199; Indels 96; Gaps 16;

40 KKQAVK-----RHHKHNLRH-----RYEFLTLGKG 66
D 10 KRRIVKVPDGDHVTNHRKNSSSHARSTQSGMSSRAARRNDQVHVGVKYLKLITIGKG 69
QY 67 TYGVKKARE-SSGRLVAIKSIRKDKIDEODLHRRREIBMSLNPHIATIHVEFEN 125
D 70 NPAKTKLAKHVTIGHEVAIKIKIDKAL-NPSSLQKLFREVKIMKQLDHNIYKLVQWMET 128
QY 126 SKKIVIVMEYASGDLVYISERPRLSERDARHFFRQIVSALHYCHONGIYVRDKLENT 185
D 129 EGTLLVLEVYASGGEVFPYLVHNGHMKKEKARAFRQIVSAVOYLHSHKNIIRDLKAENL 188
QY 186 LLDANGNKIDPFGSNLYHKGKFLQTPCGSPLYASPEIYNGKRPVYGVGVDSWISGLVLY 245
D 189 LLDQMNKIDPFGSNFSLGNKLDTPCGSPPYAPBELFSGKTYDGEPEVDWISGLVLY 248
QY 246 ILVHGTWPFQGDHKTIVKQISNGAVREPPKPSDAC-GLIRWLIMVNPFRATLEDVAVSH 304
D 249 TLVSSSLFQDGNLKELEBRVLRGKTRIPFVWSTDCENILKFLVIYNPQRSLSLNIMKD 308
QY 305 WVVWVGYTTTGVGEQALREGGHPSD-----FGASMDWLRSSRPL 347
D 309 RMMNVWGY-----EDDELKPFLEPPDQIDECRIEKLQIFOLGFKKALILEVEKE--- 359
QY 348 LENGAKVCSFFKOHVPGGGSTVPGLEROHSLKKSRKENDMAONLQGDPAEDTSSRPGASS 407
```

```
D 360 -----KREDIATYLLG-----ERKSDMDAS--BITMAOSLSHSSINVSSSLQGH- 404
QY 408 LKLPRKGLKKK---SSTSSEVQEDPQELRPVPTPGQPPVAVSLPRKGLKKSRORS 464
D 405 ---PAGVITREHVTSSASGSSASPSRYSRSATATGASITAGSLASANAQHQOQSSA 461
QY 465 GYSSPPEPSESSEGLLDASDVVSGDPVEQKSPQASGLLIH-----RKGIKLNGKFSRTA 519
D 462 APSSGSSSSSRSSQNDAA-ATAAGGVWMSGTRHGQVQMRAPTSRQATISLQPPSYXP 520
QY 520 LEGTT-----PSTFGSLDQGLASSHPAARPS 544
D 521 SSNTTQIAQIPPLFNRSSTATSS--AAQPS 548
```

RESULT 7

```
T13741
hypothetical protein 22E5.8 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13741
R:Murphy, L.; Harris, D.; Barrell, B.
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1398 <MUR>
A:Cross-references: UNIPROT:077268; EMBL:AL031765; NID:e1371523; PID:e1329905; PIDN:CAA2
C:Gene:
A:Cross-references: FlyBase:FBgn0000667
A:Interons: 205/3; 227/1; 322/3; 688/3; 782/3; 814/2; 1363/3
A>Note: EG:22E5.8
```

```
Query Match 20.2%; Score 664; DB 2; Length 1398;
Best Local Similarity 31.1%; Pred. No. 1.5e-19;
Matches 210; Conservative 92; Mismatches 216; Indels 158; Gaps 25;

10 PQQASASALSESARPLADGLIKSPKPLMKQAVKHHNKNLHRYEFLTLGKPTVG 69
D 110 PGPSPSSAVG-----AGGI--SGKDLKKEPMRVGF-----YDIERTIGKUNA 153
QY 70 KYVKARER-SGRLVAIKSIRKDKIDEODLHRRREIBMSLNPHIATIHVEFENSCK 128
D 154 VKKLARHRTKNEVAIKIKIDSGL-DQNTLQKVREIVIMKRLKPHIILKLVQWETKGM 212
QY 129 IYIVMEYASGDLVYISERPRLSERDARHFFRQIVSALHYCHONGIYVRDKLENTILD 188
D 213 IYIVSEVYASGGEIIPYIAKYGMSSSARFQWQIISAVEYCHKKKGIYHRDLKAENLILD 272
QY 189 ANGNKIDPFGSNLYHKGKFLQTPCGSPLYASPEIYNGKRPVYGVGVDSWISGLVLY 248
D 273 LMMNKIDPFGSNFSGNFKGELATWCGSPPYAPBELFSGKTYDGEPEVDWISGLVLY 332
QY 249 HGTWPFQGDHKTIVKQISNGAVREPPKPSDAC-GLIRWLIMVNPFRATLEDVAVSHW- 306
D 333 CGALPFDSTQISLDRVLRSGRFRIPFMSBSCCHILIRMLVLBTRRTYTDQIRHNM 392
QY 307 -----VNMGYTTTGVGEQALREGGHPSDGFASMDW- 340
D 393 CPLELHVLIAKYNLGAARQTSV-----EPSEDLIRI-NAEYVIGSGDTRASLKKNTYDH 447
QY 341 -----RRSSRPILLENG-----AKVCSFFKOHVPG- 366
D 448 VAAIYLLIODRVSHKKEQSNGLASALASSTYSASRMITSRRNDHPQOQSQOQSKTIST 507
QY 367 STVPGLEROH-----LKSRKENDMAONLQGDPAEDTSSRPGKSLKP-KGIILK 417
D 508 SSIILAKQCHRLSHNQVLMSENAHNGAIPYVPBPBGYAKKGPQLPLPLTGHSHL 567
QY 418 KSTTSGEVQED-----PQELRPVPTPGQVPA-VSLLPKGLIKKSRQRESGYSS 469
```

Db 568 TGYLNGGVGVNDASGILPLMRTPLP-TPASPAFNSCSTSRVG-----RHSLSSSS 619

Qy 470 P-----EPSESG-----ELLDASDVFGSDPEQKSPQASGLLHR---KGI 508

Db 620 PSHHPVAISLISLNDNPSLANRCRHMMA-----CGGVGVAVGVPLASKQHLQITSEFI 674

Qy 509 LKLNKFSRTALEGTT-----DSTGSLDQLASSHPARSPS-----GA 549

Db 675 IQGSTEDCFLALIQGSAVAAGKDDPPKASSVGVGPVPASTTPTSTAGPESGAPCPGE 734

Qy 550 VSEDSI---LSESPD 562

Db 735 INKXIKTKWSSSSSPD 750

RESULT 8

T52633
serine/threonine-specific protein kinase (EC 2.7.1.1-) AKIN11 [validated] - Arabidopsis t
N/Alternate names: SNF1 protein kinase omolog AKIN11
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #ext_change 09-Jul-2004
C/Accession: T52633
R/Bhalero, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machid
Proc. Natl. Acad. Sci. U.S.A. 96, 5522-7, 1999
A/Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein ki
A/Reference number: Z25116; MUID:99238528; PMID:10220464
A/Accession: T52633
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-512 <BHA>
A/Cross-references: UNIPROT:P92958; EMBL:X99279; PIDN:CAA67671.1
A/Experimental source: cultivar Columbia
C/Genetics:
A/Gene: AKIN11
C/Function:
A/Description: EC 2.7.1.1-; serine/threonine-specific protein kinase AKIN11 [validated, M
Complements SNF1 mutations in yeast
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase

Query Match 19.7%; Score 649.5; DB 2; Length 512;
Best Local Similarity 38.6%; Pred. No. 2.2e-19;
Matches 152; Conservative 65; Mismatches 126; Indels 51; Gaps 9;

Qy 57 YEFLETLGKGYGKVKARE-SSGRVATKSIKDKIKDEQDLHRRREIEMSSLNPH 115

Db 20 YLQKTLGIGSGFGKXIAHVVTGKVAIKIINRRIKMMEMEKVRRRIKILRLFMHP 79

Qy 116 IIAHVEPENSRIYIVMEYASRGDIYDISSRPILSERDARHFFROIYSALHYCHONGI 175

Db 80 IIRQYEVIEITTDIYVMEYVSGELFDYIVKEGRQDEBANKFPQOIISSGYEYCHRMNV 139

Qy 176 VHRDLKLENILLDANGNIKIDFGLSNLYHKGKFLQTFGSGPLVYASPEIIVNGKPYVPEV 235

Db 140 VHRDLKLENILLDSCNLIKIDFGLSNVARDGHFLTKTSCSPVYAPAVISGKLYAGPBY 199

Qy 226 DMSVIGVLLYLIVHGMPPDGDHKTLYVQISNGAVREPK-PDSACGILRMILMNPFR 294

Db 200 DWSGCVILLYALLCGLPDPDENIPNLFYKIKGGIYTLPSHSSEKARDIIPRLIYDPVK 259

Qy 295 RATLEDVASHVMVWNGYTTGVEGEALREGGHPSGDPRASMAADMLRSGRPLENGAVY 354

Db 260 RTIIEIRORHF-----QTHLPRIYAVPPTDVEAKKI 294

Qy 355 CSFFQHPVPGSGSTVGLERQSLK--KSRKENDMAQNLQGPADDTSSRPGKSLKLPK 412

Db 295 NEEIYQEVYV-----MGFDRNQVLESLRRRTQND-----ATVYLLLDNRRVPS 340

Qy 413 GILKK--KSSSTSGVEQDEPQELRPVDPPIPGPV 444

Db 341 GYLSSEFQETTDG---SNPMR--TPBAGASPV 368

RESULT 9

A26030
serine/threonine-specific protein kinase (EC 2.7.1.1-) SNF1 - yeast (Saccharomyces cerev
N/Alternate names: protein YBR477W
C/Species: Saccharomyces cerevisiae
C/Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #ext_change 09-Jul-2004
C/Accession: A26030; S69644
R/Celenza, J.L.; Carlson, M.
Science 233, 1175-1180, 1986
A/Title: A yeast gene that is essential for release from glucose repression encodes a p
A/Reference number: A26030; MUID:86289463; PMID:3526554
A/Accession: A26030
A/Molecule type: DNA
A/Residues: 1-633 <CEL>
A/Cross-references: UNIPROT:P06782; EMBL:M13971; NID:g172629; PIDN:AAA5058.1; PID:g1726
R/Dierich, F.S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A/Reference number: S69554
A/Accession: S69644
A/Molecule type: DNA
A/Residues: 1-633 <DIE>
A/Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64904.1; PID:g927732; GSPDB:GN000
C/Genetics:
A/Gene: SGD:SNF1; MIPS:YBR477W
A/Cross-references: SGD:S0002885; MIPS:YBR477W
A/Map position: 4R
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threoni
A/Note: required for expression of glucose-repressed genes in response to glucose depriv
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransf
F:51-69/Region: protein kinase ATP-binding motif
F:84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted
F:182,186/Binding site: magnesium (Asn, Asp) #status predicted
F:210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predict.

Query Match 19.7%; Score 648.5; DB 1; Length 633;
Best Local Similarity 33.1%; Pred. No. 2.9e-19;
Matches 161; Conservative 79; Mismatches 133; Indels 113; Gaps 14;

Qy 47 HHHKINLRH-----RYEFLTLGKGYGKVKKA-RESSGRVY 82

Db 22 HHHHHHHHGGGSSSTLNPPKSLADGHNIGNYQVKTGLGGSPFKVLAHYTTTGGKY 81

Qy 83 AIKSRKDKIKDEQDL-HIRREIEMSSLNPHITIAHVEPENSRIYIVMEYASRGDI 141

Db 82 ALKIKNK-KVLAKSDMQGRIEREIEIYLRILRHPHIILKYDVATKSKDEIIMVIEYAG-NEL 139

Qy 142 VYISERPRLSRDARHPRROIYSALHYCHONGIYHRDLKLENILLDANGNIKIDFGIS 201

Db 140 FDIYIQORDMSQEARFRFQOIIISAVEYCHRIKRIYHRDLKPNLLDDELNVKIDFGIS 199

Qy 202 NLHYHKGKFLQTFGSGPLVYASPEIIVNGKPYVGEVDSWSGLVLLYLIVHGMPPDGDHKT 261

Db 200 NIMTDGNFLTKSCGSPVYAPAVISGKLYAGVEVDWMSGVILLYWLCKRLFPDDSEIYV 259

Qy 262 LVQISNGAYRPPKPS-DACGILRMILMNPTRATLEDVASHVMVWNGYTTGVEGEQA 320

Db 260 LFKNISNGVYTLPKFLSPAGLIKRMILVNLIRLSIHEIMQD----- 303

Qy 321 LREGGHPSGDPRASMAADMLRSGRPLENGAKVCSFFQHPVPGSGSTVGLERQSLK 380

Db 304 -----DW-----FVVDLP-----EVLLEP 317

Qy 381 SRKENDMAQNLQGPADDTSSRPGKSLKPKGILKKSSSTSGVEQDEPQELRPVDPPI 440

Db 318 DLKHPHEENENNDDKKQSS---PDNDEIDNIVNIISSITNGYKDELDYESLESSEDP 374

Qy 441 --GQVPVAVSL-PRKGIKKSRQRESGYSSPPESEGLLDASDVFGVSGDP---VQ 493

Db 375 AFNEIRDAVMIKENKSLIKDKMKANS-----VSDDELDTFLSGSPPTPQOOS 421

QY 494 KSPQAS 499
DB 422 KSHQKS 427

RESULT 10

gik protein - chicken
JC7500
N.Alternate names: Qln-induced kinase
C.Species: Gallus gallus (chicken)
C.Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C.Accession: JC7500
R.Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.
Biochem. Biophys. Res. Commun. 276: 564-570, 2000
A.Title: The new serine-threonine kinase, Qik, is a target of the qin oncogene.
A.Reference number: JC7500
A.Contents: Embryo fibroblasts
A.Accession: JC7500
A.Molecule type: mRNA
A.Residues: 1-798 <Xia>
A.Cross-references: UNIPROT:Q91A88; GB:AF219232
C.Comment: This protein, a member of the AMPK/SNFI family of serine/threonine kinases,
C.Genetics:
A.Gene: qik
C.Keywords: protein kinase

Query Match 19.7%; Score 648; DB 2; Length 798;
Best Local Similarity 29.4%; Pred. No. 3.8e-19;
Matches 195; Conservative 94; Mismatches 223; Indels 152; Gaps 17;

QY 11 SOAPASALASASAPLADGLIKSPKPLMKQAVKHHKHLRHHYEFLETLGKGYTK 70
DB 5 SEDASVPAPSAQPPRLRVGF-----YDERTLGKGNFV 39
QY 71 VKKARESSGRV-VAIKSIRKDKIODEODLHTRREIEIMSNPHI1AHEVFENSKXI 129
DB 40 VKLAHRYATKQVAKIIDKTRL-DPSNLEKTYREVIQKLNHPIHIIKLYGMERKML 98
QY 130 VIWMEYASRGDIYIISERPLSERDARHFPRQIYSALHYCHONGIVHDLKLENTLDA 189
DB 99 YIVTFEPAKNGEMFDHILTSNGHLSSEBARCKFWQIISAVGYCHSHHIVHDLKTEMLLDA 158
QY 190 NGNITADFGSLNLAHKKFLQTFGSGPLYASPELVNGKPYGPEVDSLSGLVLYLVH 249
DB 159 NNNTIKLADFGKGNFYKSGEPLSTWCGSPYAAPEVGEKYGPHLDIWSLGVLYLVLC 218
QY 250 GTMPFGODHKTLYVQISNGAYREPKPSDAC-GILRWLWNPTRATLEDVASHMWNV 308
DB 219 GSLPFDGNLPTLRQVLEGRFRIRYFMSDECEITLRLVVDPTKRITISQIKQHKM- 277
QY 309 WGYTTGVEQEARLREGHPSGDGFRASMAADWLRRSSRPL--LENGAKVCSFKQHPVG 365
DB 278 -----QADPSLRQOQSLSFQMNVNSNLGYNQVLGI 310
QY 366 GSTVGLEROSIKK-----SRKENDMAQLQODPAEDTSSRPGKSL 408
DB 311 MQLT-GIDRQRTVESLQNSSTNHPAIIYLLERLKEYSLSLSPRATGROQRSSSEI 369
QY 409 ---KLPRKILKKKSSSTSGEVEDPQELRPV-----PDPGQPV-----PAVSLPR 452
DB 370 SNAEMFQSLSTSETLRSSILYQPOSILQPSQAMDCMNNPLQDPVFPVDPNNGGLFR 429
QY 453 KGLIKSFQRESGYSSPEPSESGLLDASDVFGSDPYEOKSPQASGILLRKGITLKN 512
DB 430 NRSISPSLSLETTI--SEVRQEKELD--EIKAVDHPRIKIS-----N 469
QY 513 GKFEFTALEGTTPTSFGLDQLASHPARPRSGAVSDESLTSSSEFTQ----- 563
DB 470 TSRHHTLAETVTHFYQHAAPCIVISSAS---PTEGTSDBCLTSSSSNDSSVALSSCIA 525
QY 564 -----LDLPERLPEPLRGCVSVDNL-----GLGQPSSEGIKR 597

DB 526 GQWMTGSPATAMTSAFLASQSDAPVLQYQCGMGASLLPVFQEGRRASDTSITQGLKA 585
QY 598 WMOE 601
DB 586 FROQ 589

RESULT 11

probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
N.Alternate names: SNF1-related protein kinase
C.Species: Cucumis sativus (cucumber)
C.Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C.Accession: T10449
R.Stumpel, N.J.
submitted to the EMBL Data Library, December 1996
A.Reference number: T17020
A.Accession: T10449
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-504 <GUN>
A.Cross-references: UNIPROT:P93113; EMBL:Y10036
A.Experimental source: cv. Masterpiece; cotyledon
C.Function:
A.Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C.Superfamily: AMP-activated protein kinase; protein kinase homology
C.Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F.6-260/Domain: protein kinase homology <KIN>

Query Match 19.5%; Score 641; DB 2; Length 504;
Best Local Similarity 50.4%; Pred. No. 4.7e-19;
Matches 127; Conservative 47; Mismatches 76; Indels 2; Gaps 2;

QY 57 YEFLETLGKGYGVKAKAES-SGRVLAIKSIRKDKIODEODLHTRREIEIMSNPH 115
DB 8 YKLGKTLGSGRGKVIKHAHALTGKVAIKILNRKIKNLDWEKVRREIKILRFMPRH 67
QY 116 IIAIHEVENSCKIYIWEYASRGDIYIISERPLSERDARHFPRQIYSALHYCHONGI 175
DB 68 IIRLYEVEITPSDIYVWEYVSGELFPYIYEKGLQDEANFPQOIIISGYEYCHRMV 127
QY 176 VHRDLKLENTLDDANGIKIDPGLSNLYHKKFLQTFGSGPLVYASPELVNGKPYGPEV 235
DB 128 VHRDLKPEVLDLSDSCNKKIADFGLSNTRDHFLLTKTSGSNVYAPETYSKLYAGPEV 187
QY 236 DWSLGLVLLYLVHGTMPFGODHKTLYVQISNGAYREPKP-PSDACGILRWLWNPTR 294
DB 188 DWSGCVILLYALLCGTLFPDDENIPNLFPKIKGIGITLPSHLSGARELIPMLVVDPMK 247
QY 295 RATLEDVASHMW 306
DB 248 RITPEIRQHPW 259

RESULT 12

UC1446
serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N.Alternate names: protein kinase SNF1 homolog
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C.Accession: UC1446; S58266; S66334
R.Luegten, U.; Thomas, M.; Blanchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A.Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A.Reference number: UC1446; MUD:99013041; PMID:1339373
A.Accession: UC1446
A.Molecule type: DNA
A.Residues: 1-512 <LEG>
A.Cross-references: UNIPROT:Q38997; GB:M91023; NID:9166599; PIDN:AAA2736.1; PID:9166600
R.Thiemmer, F.; Kirchner, M.; Teuber, R.; Dietrich, P.
submitted to the EMBL Data Library, May 1995
A.Description: Differential accumulation of the transcripts of 22 novel protein kinase g

Query Match	19.4%	Score 639.5;	DB 1;	Length 512;
Best Local Similarity	35.8%	Pred. No. 5.5e-19;		
Matches 156;	Conservative 72;	Mismatches 139;	Indels 69;	Gaps 11.

RESULT 13

A7Title: Characterization of tobacco protein kinase NPK5, a homolog of *Saccharomyces cerevisiae* Npk1p.

Query Match	19.2%;	Score 631;	DB 1;	Length 511;
Best Local Similarity	49.6%;	Pred. No. 1.2e-18;		
Matches 125;	Conservative 48;	Mismatches 77;	Indels 2;	Gaps 2

RESULT 14

A/Molecule type: LIGAND
A/Residues: 1-179, 'T', 181-270, 'G', 272-402, 'RO', 405-

A:Description: catalyzes the formation of peptidyl-L-serine-phosphate or peptidyl-L-threonine
A:Note: in vivo substrates include hydroxymethylglutaryl-CoA reductase (NADPH) and acetyl

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 11, 2005, 14:13:49 ; Search time 44.2557 Seconds
(without alignments)
1064.351 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293

Sequence: 1 MESVALLQRPQAPSASALA.....DCQEVYAAVRAALGICSKLS 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/AA_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	4	US-09-579-664B-11 Sequence 11, App1
2	3293	100.0	631	4	US-10-355-975A-11 Sequence 11, App1
3	3278.5	99.6	630	4	US-10-355-975A-38 Sequence 38, App1
4	703.5	21.4	729	2	US-08-677-298-2 Sequence 2, App1
5	703.5	21.4	729	4	US-09-523-849-33 Sequence 33, App1
6	700.5	21.3	602	4	US-09-949-016-7417 Sequence 7417, App1
7	700.5	21.3	602	4	US-09-949-016-7418 Sequence 7418, App1
8	698.5	21.2	724	4	US-09-949-016-8234 Sequence 8234, App1
9	694.5	21.1	713	4	US-09-538-092-1022 Sequence 1022, App1
10	694.5	21.1	713	4	US-09-949-016-6214 Sequence 6214, App1
11	691.5	21.0	793	4	US-09-523-849-32 Sequence 32, App1
12	688	20.9	745	4	US-09-523-849-36 Sequence 36, App1
13	680	20.6	724	4	US-09-984-890-2 Sequence 2, App1
14	680	20.6	724	4	US-10-274-194-2 Sequence 2, App1
15	678.5	20.6	691	4	US-09-949-016-8255 Sequence 8255, App1
16	678.5	20.6	691	4	US-09-949-016-8256 Sequence 8256, App1
17	677.5	20.6	722	4	US-09-984-890-4 Sequence 4, App1
18	677.5	20.6	722	4	US-10-274-194-4 Sequence 4, App1
19	675	20.5	779	4	US-08-817-832B-31 Sequence 31, App1
20	673.5	20.5	722	4	US-08-817-832B-32 Sequence 32, App1
21	653	19.8	604	4	US-09-523-849-35 Sequence 35, App1
22	648.5	19.7	633	3	US-08-557-006C-43 Sequence 43, App1
23	648.5	19.7	633	4	US-09-538-092-212 Sequence 212, App1
24	648.5	19.7	633	4	US-09-633-328B-3 Sequence 3, App1
25	648.5	19.7	633	4	US-09-824-735-3 Sequence 5, App1
26	648.5	19.7	1203	4	US-09-799-875-5 Sequence 2, App1
27	645.5	19.6	512	4	US-09-633-328B-2 Sequence 2, App1

28	642	19.5	557	4	US-09-949-016-10174 Sequence 10174, App1
29	636.5	19.3	552	3	US-09-824-735-4 Sequence 4, App1
30	635.5	19.3	345	3	US-09-101-146-1 Sequence 1, App1
31	635.5	19.3	776	4	US-09-523-849-34 Sequence 34, App1
32	631	19.2	511	4	US-09-633-328B-4 Sequence 4, App1
33	628.5	19.1	552	4	US-09-538-092-1212 Sequence 1212, App1
34	625.5	19.0	418	4	US-09-248-796A-18441 Sequence 18441, App1
35	625	19.0	257	3	US-09-101-146-6 Sequence 6, App1
36	623.5	18.9	552	3	US-08-557-006C-40 Sequence 40, App1
37	619.5	18.8	257	2	US-07-857-224B-25 Sequence 25, App1
38	613.5	18.6	778	4	US-10-116-326-2 Sequence 2, App1
39	613.5	18.6	778	4	US-10-003-690-2 Sequence 2, App1
40	599.5	18.2	668	4	US-09-930-181-2 Sequence 2, App1
41	589.5	17.9	521	4	US-09-949-016-7570 Sequence 7570, App1
42	549	16.7	1064	4	US-09-538-092-154 Sequence 154, App1
43	547	16.6	703	4	US-10-116-326-6 Sequence 6, App1
44	528.5	15.0	603	4	US-09-930-181-17 Sequence 17, App1
45	525	15.9	260	2	US-07-857-224B-28 Sequence 28, App1

ALIGNMENTS

RESULT 1

US-09-579-664B-11

Sequence 11, Application US/09579664B

Patent No. 6514719

GENERAL INFORMATION:

APPLICANT: Immunex Corporation

APPLICANT: Bird, Timothy A.

APPLICANT: Virca, G. Duke

APPLICANT: Martin, Unja

APPLICANT: Anderson, Dirk M.

TITLE OR INVENTION: NOVEL MURINE AND HUMAN KINASES

FILE REFERENCE: 2923-A

CURRENT APPLICATION NUMBER: US/09/579,664B

CURRENT FILING DATE: 2000-05-26

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn version 3.1

SEQ ID NO 11

LENGTH: 631

TYPE: PRT

ORGANISM: Mus musculus

US-09-579-664B-11

Query Match	Score 3293; DB 4;	Length 631;
Best Local Similarity 100.0%;	Pred. No. 2.4e-277;	Indels 0; Gaps 0;
Matches 631; Conservative	0; Mismatches	0; Indels
QY	1	MESVALLQRPQAPSASALASARPLADGLTSPKPLMKQAVKHHKHLRHRVEFL 60
DB	1	MESVALLQRPQAPSASALASARPLADGLTSPKPLMKQAVKHHKHLRHRVEFL 60
QY	61	ETLGGKTYGVKKARSSGRVLAIKSIRKDKDEBDLHIREIEMSLNPHIAT 120
DB	61	ETLGGKTYGVKKARSSGRVLAIKSIRKDKDEBDLHIREIEMSLNPHIAT 120
QY	121	EYFENSSKVIYWEVSRDLYDISRRPLSRDRAHFRQVSLAHYCHONGIYHRL 180
DB	121	EYFENSSKVIYWEVSRDLYDISRRPLSRDRAHFRQVSLAHYCHONGIYHRL 180
QY	122	EYFENSSKVIYWEVSRDLYDISRRPLSRDRAHFRQVSLAHYCHONGIYHRL 180
DB	122	EYFENSSKVIYWEVSRDLYDISRRPLSRDRAHFRQVSLAHYCHONGIYHRL 180
QY	181	KLENTLLDNGNINKIADFGLSNLYHKGFLQPFQCSFLVASPIYGVKPYGEVDSWSL 240
DB	181	KLENTLLDNGNINKIADFGLSNLYHKGFLQPFQCSFLVASPIYGVKPYGEVDSWSL 240
QY	241	GVLLYTLVGTMPFDQDHTLVKQISNGAYRPPKPSDAGLIRWLMVNPTRATLTD 300
DB	241	GVLLYTLVGTMPFDQDHTLVKQISNGAYRPPKPSDAGLIRWLMVNPTRATLTD 300
QY	301	VASHMVMVNGYTTGTGVEQEARLREGRHPSGDFGASVADWLRRSSPLLENAGAVCSFFQ 360
DB	301	VASHMVMVNGYTTGTGVEQEARLREGRHPSGDFGASVADWLRRSSPLLENAGAVCSFFQ 360

QY	361	HVPGGGSYTPGLERHSLKKSRKENDMAONLOQDPADETSRRGKSIXLPGKILLKXSS	420
Db	361	HVPGGGSYTPGLERHSLKKSRKENDMAONLOQDPADETSRRGKSIXLPGKILLKXSS	420
QY	421	TSSGVEQEDPQELRPVPDTPGGPVPVAVSLPRKGIILKKSRKORESGYSSSPBPSGSEGLD	480
Db	421	TSSGVEQEDPQELRPVPDTPGGPVPVAVSLPRKGIILKKSRKORESGYSSSPBPSGSEGLD	480
QY	481	ASDVAVSGDPVYQKSPQASGLILHRGIIKLNKKEPRTLEGTPSTFGSLDLSAHPA	540
Db	481	ASDVAVSGDPVYQKSPQASGLILHRGIIKLNKKEPRTLEGTPSTFGSLDLSAHPA	540
QY	541	ARPSRPSGAVSDEDLSLSSFPQDLPERLPEPLPLGCVSVNDLRGLGEPSPSGLRWMQ	600
Db	541	ARPSRPSGAVSDEDLSLSSFPQDLPERLPEPLPLGCVSVNDLRGLGEPSPSGLRWMQ	600
QY	601	ESLGGSCFSLTDCQEVTAAYRQALGICXSLKS	631
Db	601	ESLGGSCFSLTDCQEVTAAYRQALGICXSLKS	631

```

RESULT 2
US-10-355-975A-11
: Sequence 11, Application US/10355975A
: Patent No. 6759223
: GENERAL INFORMATION:
: APPLICANT: Immunex Corporation
: APPLICANT: Bird, Timothy A.
: APPLICANT: Virca, G. Duke
: APPLICANT: Martin, Unja
: APPLICANT: Anderson, Dirk M.
: TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
: FILE REFERENCE: 2923-B
: CURRENT APPLICATION NUMBER: US/10/355, 975A
: CURRENT FILING DATE: 2003-01-30
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 11
: LENGTH: 631
: TYPE: PRT
: ORGANISM: Mus musculus
: US-10-355-975A-11

```

Query Match	100.0%;	Score 3293;	DB 4;	Length 631;
Best Local Similarity	100.0%;	Pred. No. 2,4e-277;		
Matches 631;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MESVALLORPSQAPASALASESPAPLADGLIKSPKPLMKQAVVRHHKHLRRHYEL	60		
Db 1	MESVALLORPSQAPASALASESPAPLADGLIKSPKPLMKQAVVRHHKHLRRHYEL	60		
QY 61	ETLKGTYGKVKAKARESSGRLVAIKSRDKDKEDODLHIREIEMSLNPHI1AIH	120		
Db 61	ETLKGTYGKVKAKARESSGRLVAIKSRDKDKEDODLHIREIEMSLNPHI1AIH	120		
QY 121	EVFENSXIIVLVMYASRGDLVDYISERPRLSERDARHPFROIYSALHYCHONG1VHRDL	180		
Db 121	EVFENSXIIVLVMYASRGDLVDYISERPRLSERDARHPFROIYSALHYCHONG1VHRDL	180		
QY 181	KLENIILDANINIKIADRGLSNLVHKRGFLQFCCSPLYASEIYNGKPYGPEVDSMSL	240		
Db 181	KLENIILDANINIKIADRGLSNLVHKRGFLQFCCSPLYASEIYNGKPYGPEVDSMSL	240		
QY 241	GVLLYIILVHGTMPEFGDQDKTLVQKISNGAYREPKSPSDACGLIRMLVMVNTFRATLED	300		
Db 241	GVLLYIILVHGTMPEFGDQDKTLVQKISNGAYREPKSPSDACGLIRMLVMVNTFRATLED	300		
QY 301	VASHTVMVNWGTYTGVGEQBALREGHPSGDFGRASMDLRRSSRLLENGAKVCSFPKQ	360		
Db 301	VASHTVMVNWGTYTGVGEQBALREGHPSGDFGRASMDLRRSSRLLENGAKVCSFPKQ	360		
QY 361	HVPGGGSIVPLELEQHSILKSRKENDMAQNTLQGDPAETSSRPKASSLKLPGKIIKKTSS	420		
Db 361	HVPGGGSIVPLELEQHSILKSRKENDMAQNTLQGDPAETSSRPKASSLKLPGKIIKKTSS	420		

Db	361	IVPGGGGSTYVPLGERHSLKKRKNEMDNQLODPAEDTSSRPGKSLPLPKILKKKS	420
QY	421	TSSGEVQEDPQELRPVPDPPTGQVPDAVSLLPKKGILKKSRQRESGYSSPEPSEGLLD	480
Db	421	TSSGEVQEDPQELRPVPDPPTGQVPDAVSLLPKKGILKKSRQRESGYSSPEPSEGLLD	480
QY	481	ASDVAVSGDPVEQKSPQASGLLIHKKGILKXNGKFSRTLEKTPETPSLDDLAASHPA	540
Db	481	ASDVAVSGDPVEQKSPQASGLLIHKKGILKXNGKFSRTLEKTPETPSLDDLAASHPA	540
QY	541	ARPSRPSGAVSEDSILSSSPDQLDLPERLPEPTPLRGCVVNINGLLEQPPSEGLRRWQ	600
Db	541	ARPSRPSGAVSEDSILSSSPDQLDLPERLPEPTPLRGCVVNINGLLEQPPSEGLRRWQ	600
QY	601	ESLSPDSCTSLDDQEVTAIYRQALGICSTLS	631
Db	601	ESLSPDSCTSLDDQEVTAIYRQALGICSTLS	631

```

RESULT 3
US-10-355-975A-38
; Sequence 38. Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355,975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975A-38

```

Query	March	99.6%	Score 3278.5	DB 4	Length 630
Best	Local Similarity	99.8%	Pred. No. 4.4e-276		
Matches	630	Conservative	0	Mismatches	0
				Indels	1
				Gaps	1
QY	1	MESVALLQRPQASASALASESARPLADGLIKSPKPLMKQAVVRHHKHNLRAREYFL	60		
Db	1	MESVALLQRPQASASALASESARPLADGLIKSPKPLMKQAVVRHHKHNLRAREYFL	60		
QY	61	ETLKGKTYGKKKARSSGRLVAIKSRKDKIKOBODLIHRRRIETMSLNHPITIAIH	120		
Db	61	ETLKGKTYGKKKARSSGRLVAIKSRKDKIKOBODLIHRRRIETMSLNHPITIAIH	120		
QY	121	EVFENSKIIVMEYASRGDLDYIYSERPRSESDARHFPFOIYSALHYCHONGIVHADD	180		
Db	121	EVFENSKIIVMEYASRGDLDYIYSERPRSESDARHFPFOIYSALHYCHONGIVHADD	180		
QY	181	KLENIILLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYVGNKRPYGVPEVDSNLT	240		
Db	181	KLENIILLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYVGNKRPYGVPEVDSNLT	240		
QY	241	GVLVLYIIVHGMPPEDGODHKTIVQIINGAAREPKPSDAGLIRMLLMNPTRRATIED	300		
Db	241	GVLVLYIIVHGMPPEDGODHKTIVQIINGAAREPKPSDAGLIRMLLMNPTRRATIED	300		
QY	301	VASHMVMVMGTYTIVGEQALREGGHSPDGFGRASMADWLRSRPLLENAAKYSFPKQ	360		
Db	301	VASHMVMVMGTYTIVGEQALREGGHSPDGFGRASMADWLRSRPLLENAAKYSFPKQ	360		
QY	361	HVPGGGSITVRLERQHSIKRSRKNDMAQNIQDPADBTSSRPKSSIKLPEKILUKKSS	420		
Db	361	HVPGGGSITVRLERQHSIKRSRKNDMAQNIQDPADBTSSRPKSSIKLPEKILUKKSS	420		


```
Db      235 DWSVLGVILYTLVSGSLPFGCGNMLKRLERVLGKTRIPFYMSTDCENLKRFLVLANPIK 294
Qy      295 RATLELVASHMWNWNGYTTTGVGEQALREGGHPSGDFGRASMDWLRSSRPILLENGAKY 354
Db      295 RGLLEIMKDRMINAGH-----BEDELKPFVEPELIDSDQKRIDIM-----Y 336
Qy      355 CSPEFKHVPGGGTVGTLERQSLKSKRKENMDMAQNL---QGDPAETSSRPKSSILKLP 411
Db      337 GMYGSOB-----EIQESLSKMKYDEITATYLLLRKSESLDASSSSSSNLSTA 395
Qy      412 KGLIKKSSSTSGE---VQ---BDPELRPVDPFGQPVPAVSLPRK-----GIL 456
Db      386 KVRPSSDLNNSTQSPHMKVQRVSSSQKQRRYSDDHGAIPESVVAPEKSSQSTADGL 445
Qy      457 KR---SRQES-----GYSSPSPSESGELLDASDVFSVSGDPVEQSPQ 497
Db      446 KEGGISRSKSSGSAVGKGIAPASPMIGNASNENKADIPERKKSSTV-----PSSNT 497
Qy      498 ASGLLHRK-----GILKNGKFSRTALEGTTPTFGSLDQLASH---PAA 541
Db      498 ASGGMTRRTYTCSERTYADRHSHVIO-NGKENSTIPDQRT-----VASTHSISSAA 548
Qy      542 RPSR---PSGAVSEDS 554
Db      549 TPDRIRFPRTASRST 564
```

```
RESULT 6
US-09-949-016-7417
; Sequence 7417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7417
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7417
```

```
Query Match      21.3%; Score 700.5; DB 4; Length 602;
Best Local Similarity 31.2%; Pred. No. 5e-52;
Matches 196; Conservative 100; Mismatches 216; Indels 117; Gaps 21;

Qy      23 SARPLADGLIKSPKPLMKQAVKRRHHKNNLRH-----YEFLETGKGYGVKKAR- 76
Db      34 SSRPTLGH--DSKPSKSNMTRGRNSATSADQPHIGNYRLKTTGKNFAVKLARIH 91
Qy      77 SSGRLVAISIRKDKIKDEODLHTRREIEMSLNHPHIIAIEHVEFNSSKIIVMEYA 136
Db      92 LIGKEVAVKIIDTQTL-NSSSLQKLFREVRIMKVLNHPNIVKLFVIEITEKTLIVMEYA 150
Qy      137 SRGDLVDYISERPRLSERDARHFFRQIVSALHYCHQNGIVHRDLKLENTILDANGNIKIA 196
Db      151 SGGEVFDYLVAGRMKEKEARAKFRQIVSAVOYCHQKFIHRDLKAENILDDADWNIKIA 210
Qy      197 DFGLSNLVHKGKFLQTFGCSPLVYASPEIYVNGKPYVGEVDSWSLGLVLYIVHGTMPFDG 256
Db      211 DFGSNIEFTFGNKLDTFGCSPPYAABELFGKKYDGPVDSVLSGLVILTVSGSLPFDG 270
```

```
Qy      257 QDHKTLVKQISNAGVREPPKPSDAC-GLIRMLMNVPTRATLEDVASHMWNWNGYTTGV 315
Db      271 QNKLREBVLKREKTRIPFYMSTDCENLKKPLILPSRGTLEQIMKRMNVGH----- 336
Qy      316 GEQALREGGHPSGDFGRASMDWLRSSRPILLENGAKYCSPEFKHVPGGGTVPELJERO 375
Db      327 -BEDELKPFVEPELIPVDKDP-----RRTELAVSMG-----YTRERI 361
Qy      376 HSLIKSRKENMDA-----QNLQGD-----PAED-TSSRPKSLKLPKGI----- 414
Db      362 DSIIVGQRNVEWATYLLLYGKSSSELEGDTITLKPRPSADLTNNSAPSPHKVQSVSANP 421
Qy      415 -----LKKSSSTSGE---VQEDPQGLRVPDPV---PQGPVPAVSL 450
Db      422 KORRPSDQAPAIPTSNYSSTKQTSNNANENKRPEDRESGRASSTAKVPASPLP----- 476
Qy      451 PRGILKGRORBSGGYSSPSPSESGELLDASDVFSVGDVPQKSPQASGLLHRKGIK 510
Db      477 ---GLERK-----TTPTPS-TNSVLTSTNSRNPFLERASLQASIQNGKDSLT 524
Qy      511 LNKESRTALEGTTPTFGSLDQLASHPARPSPSGAVSBSISSEFPQDLPERL 570
Db      525 MFSRSTASAGAASAAVSAARPRQHKSMASVHPNKASG-----LPPTESNCEVPRPROV 578
Qy      571 PETPLRGYVNDNLRLGEOPPESEGLKRW 599
Db      579 PP---ACPCCLPIRPHQ-----QQW 597
```

```
RESULT 7
US-09-949-016-7418
; Sequence 7418, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7418
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7418
```

```
Query Match      21.3%; Score 700.5; DB 4; Length 602;
Best Local Similarity 31.2%; Pred. No. 5e-52;
Matches 196; Conservative 100; Mismatches 216; Indels 117; Gaps 21;

Qy      23 SARPLADGLIKSPKPLMKQAVKRRHHKNNLRH-----YEFLETGKGYGVKKAR- 76
Db      34 SSRPTLGH--DSKPSKSNMTRGRNSATSADQPHIGNYRLKTTGKNFAVKLARIH 91
Qy      77 SSGRLVAISIRKDKIKDEODLHTRREIEMSLNHPHIIAIEHVEFNSSKIIVMEYA 136
Db      92 LIGKEVAVKIIDTQTL-NSSSLQKLFREVRIMKVLNHPNIVKLFVIEITEKTLIVMEYA 150
Qy      137 SRGDLVDYISERPRLSERDARHFFRQIVSALHYCHQNGIVHRDLKLENTILDANGNIKIA 196
Db      151 SGGEVFDYLVAGRMKEKEARAKFRQIVSAVOYCHQKFIHRDLKAENILDDADWNIKIA 210
Qy      197 DFGLSNLVHKGKFLQTFGCSPLVYASPEIYVNGKPYVGEVDSWSLGLVLYIVHGTMPFDG 256
Db      211 DFGSNIEFTFGNKLDTFGCSPPYAABELFGKKYDGPVDSVLSGLVILTVSGSLPFDG 270
```



```

OY  VCSFPGK-----HVPGGSTVPGJEROLSKKRENNMAONQDPREDTS 400
354 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  VSQKDEWMATYIILGRKPEPEGGSLSG-----NLCCRSRSSDLNNSTLOSPPH--- 406
                                     : : : : : : : : : : : : : : : :
OY  SRPGKSLKPKGILKKKSTSGSEVOEDPOELRPVDTPGQPV-PAVSLIPE-KGILKK 458
                                     : : : : : : : : : : : : : : : :
Db  -----LKVGRSISANQK-----RRSDHAGPEIIPAVSYTKRPQANSVE 446
                                     : : : : : : : : : : : : : : : :
OY  SRQRE-----SGYSSP--EPSSGELLASDVYVSGDPVEQKSPQA 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  SEQKEEMDKDTRRLGSTITVGSKEVTAPLVPDPKKKSAGASNNVVSGSGMTRNTYV 506
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY  SGLLHRKGILKLNKFSRTALEGTPPSFGSLDOLASSHPARP 543
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  CERSTDRYALONGRDSILTEMASASWSSTGS--TVAASGSPARP 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
US-09-523-849-36
: Sequence 36, Application US/09523849
: Patent No. 6458561
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Molteni, Angela
: APPLICANT: Magnaghi, Paola
: APPLICANT: Bobotti, Roberta
: APPLICANT: Scaccheri, Emanuela
: APPLICANT: Isaacchi, Antonella
: APPLICANT: Hodgson, Dave
: TITLE OF INVENTION: HUMAN NIM1 KINASE
: FILE REFERENCE: PC-0009 US
: CURRENT APPLICATION NUMBER: US/09/523,849
: CURRENT FILING DATE: 2000-03-13
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PERL Program
: SEQ ID NO 36
: LENGTH: 745
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36

```

Query	March	20.9%	Score	688	DB	4	Length	745	
Best	Local	Similarity	32.9%	Pred.	NO	8.5e-51			
Matches	185	Conservative	89	Med.	Matches	192	Indels	96	Gaps
QY	57	YEPLETTCKGYTYGKYKAKRE--SSGRLVAIKSIRKDKIXDEODLHRRREIEMSSINHPH	115						
Db	20	YRLKTIQGNPAKYLARHLITGSEVAVKIIDKQL--NSSSLQTLFRRVRLMKYLNHN	78						
QY	116	IIAHEFVENSSKIIVIMEYASRDLVDYISERPRISERDARHFFRQIVSALHYCHONGI	175						
Db	79	IVKLFEVIEETEKTIYLIVMEYASGSEVFDVLVAHGMIKEKAAKFRQIVSAVOYCHQKEI	138						
QY	176	VHRDLKLENIILLDANGNIKIADFGLSNLTHKKPQITPFGSPLVASPELVNKPYPVGPPEV	235						
Db	139	VHRDLKAENLLLDAAOMNIKIADFGSNSNETFENFKLIDFGSPSPYAPABELFOCKKYDGPPEV	198						
QY	236	DSWSLGYVLLYLIVHGTMFPFGQDQDHTLVQIENGAYREBPKPSDAC--GLIRMLVMNPR	294						
Db	199	DWMSLGYVLLYLVSSSLPFPDQGNLKEHREYLRGKTRIDPFYNSTDCENILKKFLLINPK	258						
QY	295	RATLEDVASHWVMVNGYTTGVGSEQALREBGHPSGDPCGRASNADWLRSSRPBLENGAKY	354						
Db	259	RGTLGQIMKDRMNVNGH-----EDDEBLKYEYVFLPDYKOP-----RATLELWMSG---	303						
QY	355	CSFFQAHYPGGGSITYPGLERQSLKKSRKENDMA-----QNLQGD-----PAED	398						
Db	304	--YTHEEL-----QSLVGQQRNEVMAVYLLIGYSSLEEGTITLKPSPSD	349						

```

Oy      399  TSSRPGKSLKLPFGI-----LKSSSSSGE---VOGDPOB 432
Db      350  LTNSSAGPSPSIKVQBSVANKQRKRRFSQAGPAIPTSNYSKKTQSNMNEKRRPEEDRES 409
Oy      433  LRPEVDT---EGQPVAVSLLPKRKGIILKKSRRSGSYSSPEPSSGELLDA5DVEVSGD 489
Db      410  GRKASSTAKVDPASPLP-----GLERKK-----TTPPS--TNSVLSTSTNRSRNS 452
Oy      490  PVEQKSPASGLLHKRKGIILKNGKFSRTALEGTTPTSPFGSLDGLASSHPAARPERPSGA 549
Db      453  PLIRERASTIGQASIONGKDSLTMPGSRASTASASAAVAPRQHOXKMSASVHPNKA5GL 512
Oy      550  VSEDSITLSSSGSFQDLDLPERLP 571
Db      513  PPTES--NCEVPRPSTAPQRPV 532

```

```

RESULT 13
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-984-890-2

```

Query Match	20.6%;	Score 680;	DB 4;	Length 724;
Best Local Similarity	32.0%;	Pred. No. 4e-50;		
Matches	201;	Conservative	99;	Mismatches 195;
				Indels 134;
				Gaps 25;
Qy	6	LLOPQSPASASALASBPAPLADGLIKSPKPLMKQAVKRIHHKHLNRH----	YEFF	60
Db	11	LNERTDQPTGLHLD-----	KPSSKSMNIGRNSATADBDQPHIGNRYLL	56
Qy	61	ETLKGTYGVKKYKARE--SSGRLYAIKIRKOKIDBDLHIREIRIEMSLNPHIIAI	119	
Db	57	KTIKGNFAKYLARHLLITGEVAVKIIDTKQL--NSSSLQKFEFVIMKYLNPYVKL	115	
Qy	120	HEVENSRSKIYIVMEVASRGDLYDISERPLSERDARHFFRQIVSALHYCHONGIYHARD	179	
Db	116	FEVIEETKTYLYLMEVASGGSEVPDYLAHGRMEKEKARAKRQVAVOYCHQKFIYHARD	175	
Qy	180	LKLENIILLDANGRIKIKADPGLSNLYHHKGFLOTFCGSPPLVASPRITVNGKPYVGEYDWSMS	239	
Db	176	LKANENILLDDADMNKIKADPFGSNEFTGNKLDIDFCGSPPAAPAPLPQGGKYTDGGEVWMS	235	
Qy	240	LGVLILYLVHGTMBPFDQGDHKLTVKQISNGAYREPPRPSPDAC--GLIRMLLMVNTPTRATL	298	
Db	236	LGVLILYLVHGTMBPFDQGDHKLTVKQISNGAYREPPRPSPDAC--GLIRMLLMVNTPTRATL	295	
Qy	299	EDVASHHWVWNGYITTCVGOEALREGGHPSGDFGRASAMADMLRRSSRPFLLENKAKVCSFF	358	
Db	296	EQIKDKRMWNVGH-----EDDELKPYVEPLPYDKDP-----	RTEILWMSWG-----YT	338
Qy	359	KQHPVGGGTVPLGERHSLKRSKENDMA-----	ONTLQGD-----PAED--TSS	401
Db	339	REEL-----	QDSLGVGQYNNVMAITLLIGTKSSLEBEDTTLTKRPRPADLITNS	386
Qy	402	RPGKSSLLKPKGI-----	LKKKSTSSGE-----VOEDPQELRPV	436
Db	387	SAPSPSHKVVORASVANKPQRRPDSQDQAPLPTNSYAKTOSNAENKRRPEDEDSGRKA	446	

QY 437 PPT---PQGPVAVSLPRKGIKKSQRESGYSSPESEGGELLDASDVFSGDPVQ 493
 Db 447 SSTAKVPASPLP-----GLERKK-----TTPPTSTNSVLSTSTN-----RSR 483
 QY 494 KSPQASGLLHRRKGIKLT---NGKFSRTALEGTTPSTFGSLDQASSHPAA-RPSRPSGA 549
 Db 484 NSP-----LLERASLGQASIQNGKDS-TAPQ-RVPVAPSANHISSSGAPDRTNPRGV 536
 QY 550 VSEDSTLSS---ESFDQDLPERL-PETP 574
 Db 537 SRSSTFHAQLRQVRDQNLPGVTPASP 565

RESULT 14

US-10-274-194-2
 ; Sequence 2, Application US/10274194
 ; Patent No. 6706511
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001306DIV
 ; CURRENT FILING DATE: 2003-12-18
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 724
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-274-194-2

Query Match 20.6%; Score 680; DB 4; Length 724;
 Best Local Similarity 32.0%; Pred. No. 4e-50;
 Matches 201; Conservative 99; Mismatches 195; Indels 134; Gaps 25;
 QY 6 LLQRSQASASALASASAPLADGLIKSPKPLMKQAVKRRHKKNLNHR-----YERL 60
 Db 11 LNERDTEQPTGLHLS-----KSSSKSNMIRGRNSATSADQPHIGNYRL 56
 QY 61 ETLGKGYGVKKARE-SSGRVLAIKSIRKDKIKDEODLHIRREIEIMSSLNHPHII 119
 Db 57 KTIQGNPAKVAKLARIHLITGKEVAVKIIDKTQL-NSSSLQQLFRVRIMKVLNHN 115
 QY 120 HEVFNSSKIVIMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGIYHD 179
 Db 116 FEVIEETKTLVIMEYASGGEVFDYLVAGRMKEKAKRAKQVSAVOYCHQKFIYHD 175
 QY 180 LKLENIILDANGNIKIADFGLSNLHYKGFLOTFCGSPLYASPELVNGKPYVGEYDWS 239
 Db 176 LKAENILLDADNMNIKIADFGFSNEFFGKLDTCGSPPYAABELFOGKKYDGPEDVWS 235
 QY 240 LGVLYIYLVHGTMPDGDHKTLLVKQISNGAYRPPKPSDAC-GLIRWLLMNPTRATL 298
 Db 236 LGVLIYIYLVSSLPFDGQNLKELRERVLRGKTRIPFYVSTGCENILKFKLLINPSRGT 295
 QY 299 EDVASHMVMVNGYTTGVGEQALREGGHPSGDFGRASMAWDLRRSSRPLLENGAKVCSFP 358
 Db 296 EQIMDRMNVGH-----EDDELKPYVEPLPDYKDP-----RRTELVMVG---YT 338
 QY 359 KOHVGGGSGTVPGLERQSLKSKRKENDMA-----QNLQGD-----PAED-TSS 401
 Db 339 REEL-----QDNLVGRYNEVMATYLLLGKSSSELEGDITTLKPPPSADLTWS 386
 QY 402 RPKGSLKLPKGI-----LKKSSSTSSGE---VOEDPQELRV 436
 Db 387 SAPSSHVKQSVSNAPKORRPSDQAGPAIPTSNYSKTKQSNNAENKRPEDRESGRA 446
 QY 437 PDT---PQGPVAVSLPRKGIKKSQRESGYSSPESEGGELLDASDVFSGDPVQ 493
 Db 447 SSTAKVPASPLP-----GLERKK-----TTPPTSTNSVLSTSTN-----RSR 483

QY 494 KSPQASGLLHRRKGIKLT---NGKFSRTALEGTTPSTFGSLDQASSHPAA-RPSRPSGA 549
 Db 484 NSP-----LLERASLGQASIQNGKDS-TAPQ-RVPVAPSANHISSSGAPDRTNPRGV 536
 QY 550 VSEDSTLSS---ESFDQDLPERL-PETP 574
 Db 537 SRSSTFHAQLRQVRDQNLPGVTPASP 565

RESULT 15

US-09-949-016-8255
 ; Sequence 8255, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8255
 ; LENGTH: 691
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-8255

Query Match 20.6%; Score 678.5; DB 4; Length 691;
 Best Local Similarity 33.9%; Pred. No. 5e-50;
 Matches 194; Conservative 89; Mismatches 175; Indels 115; Gaps 23;
 QY 57 YEFLTLGKGYGVKKARE-SSGRVLAIKSIRKDKIKDEODLHIRREIEIMSSLNHPH 115
 Db 20 YRLKTIQGNPAKVAKLARIHLITGKEVAVKIIDKTQL-NSSSLQQLFRVRIMKVLNHN 78
 QY 116 IIAHEVFNSSKIVIMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGI 175
 Db 79 IYKLEFEVIEETKTLVIMEYASGGEVFDYLVAGRMKEKAKRAKQVSAVOYCHQKFI 138
 QY 176 VHRDLKENIILDANGNIKIADFGLSNLHYKGFLOTFCGSPLYASPELVNGKPYVGEY 235
 Db 139 VHRDLKAENILLDADNMNIKIADFGFSNEFFGKLDTCGSPPYAABELFOGKKYDGPED 198
 QY 236 DWSLGLVLYIYLVHGTMPDGDHKTLLVKQISNGAYRPPKPSDAC-GLIRWLLMNPTR 294
 Db 199 DWSLGLVLYIYLVSSLPFDGQNLKELRERVLRGKTRIPFYVSTGCENILKFKLLINPS 258
 QY 295 RATEDVASHMVMVNGYTTGVGEQALREGGHPSGDFGRASMAWDLRRSSRPLLENGAKV 354
 Db 299 RCTLEQIMDRMNVGH-----EDDELKPYVEPLPDYKDP-----RRTELVMVG--- 303
 QY 355 CSFFKQHPVGGSGTVPGLERQSLKSKRKENDMA-----QNLQGD-----PAED 398
 Db 304 -YTRREI-----QDNLVGRYNEVMATYLLLGKSSSELEGDITTLKPPPSADLT 349
 QY 399 -TSSRPPKGSLLKLPKGI-----LKKSSSTSSGE---VOEDPQELRV 432
 Db 350 LTNSSAPSPSHVKQSVSNAPKORRPSDQAGPAIPTSNYSKTKQSNNAENKRPEDRES 409
 QY 433 LRVPDT---PQGPVAVSLPRKGIKKSQRESGYSSPESEGGELLDASDVFSGDPVQ 489
 Db 410 GRKASSTAKVPASPLP-----GLERKK-----TTPPTSTNSVLSTSTN----- 447
 QY 490 PVEQKSPQASGLLHRRKGIKLT---NGKFSRTALEGTTPSTFGSLDQASSHPAA-RPSR 545
 Db 448 -RSRNSP-----LLERASLGQASIQNGKDS-TAPQ-RVPVAPSANHISSSGAPDRTN 499

Qy 546 PSGAVSEDSILSS--ESPDQDLPERL-PETP 574
Db 500 PRGVSSRSTFHAGQLRQVRDQONTLYGVTPASP 532

Search completed: May 11, 2005, 14:26:06
Job time : 46.2557 secs

THIS PAGE BLANK (USP 11)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:09:18 ; Search time 165.602 Seconds

(Without alignments)
1473.690 Million cell updates/sec

Title: US-09-980-464-11

Sequence: 1 MESVALLQRPSPQASASALA.....DCQEVTAAYKALGICSKLS 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16dec04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	4 AAB50056	Aab50056 Murine ly
2	3293	100.0	652	8 ABO84756	ABO84756 Murine ca
3	3068.5	93.2	630	5 AAB19885	Aae19885 Rat SNF1/
4	2824.5	85.8	628	4 AAM93360	Aam93360 Human pol
5	2824.5	85.8	628	4 ABUS3116	Abus3116 Human cel
6	2824.5	85.8	628	5 ABP69116	Abp69116 Human pol
7	2824.5	85.8	628	5 AAU79652	Aau79652 Human pro
8	2824.5	85.8	628	7 ADF76965	Adf76965 Novel hum
9	2824.5	85.8	628	8 ADL30886	Adl30886 Human pro
10	2824.5	85.8	628	8 ADL25362	Adl25362 Human SNA
11	2824.5	85.8	628	8 ADL14161	Adl14161 Novel hum
12	2824.5	85.8	628	8 ADO20172	Ado20172 Human PRO
13	2824.5	85.8	628	8 ABO84757	ABO84757 Human can
14	2824.5	85.8	628	8 ADJ96620	Adj96620 Human cal
15	2800	85.0	629	8 AAB65632	Aab65632 Novel pro
16	2800	85.0	629	8 ADI29239	Adi29239 Human MAR
17	2720.5	82.6	594	8 AAE16266	Aae16266 Human kin
18	2590	78.7	611	6 ABP96085	Abp96085 Human pro
19	2416.5	73.4	534	4 AAB71959	Aab71959 Human TGF
20	1781.5	54.1	406	5 ABP51382	Abp51382 Human MDD
21	1781.5	54.1	406	5 ABP51482	Abp51482 Human MDD
22	1699.5	51.6	661	4 AAB07847	Aae07847 Human pro
23	1699.5	51.6	661	4 AAB07846	Aae07846 Human pro
24	1699.5	51.6	661	7 ADE38421	Ade38421 Human pro
25	1699.5	51.6	661	8 ADJ75331	Adj75331 Marker ge

26	1699.5	51.6	661	8 ADL25353	Adl25353 Human ARK
27	1699.5	51.6	661	8 ADQ19734	Adq19734 Human sof
28	1698.5	51.6	660	7 ADN95766	Adn95766 Human BEC
29	1301.5	39.5	530	4 AAE00668	Aae00668 Human pro
30	848	25.8	434	4 ABB67451	Abb67451 Drosophila
31	748	22.7	752	4 AAM93956	Aam93956 Human pol
32	748	22.7	752	4 AAE11782	Aae11782 Human kin
33	748	22.7	752	5 ABB04433	Abb04433 Human neu
34	748	22.7	752	6 ABG73794	Abg73794 Human MAR
35	748	22.7	752	8 ADL32125	Adl32125 Human pro
36	744	22.6	688	5 ABB04434	Abb04434 Human neu
37	744	22.6	689	7 ADF74129	Adf74129 Human nov
38	744	22.6	752	8 ADJ96622	Adj96622 Human cal
39	743.5	22.6	769	5 ABP62966	Abp62966 Human pol
40	739	22.4	688	6 AAB33554	Aae33554 Human mic
41	739	22.4	688	6 AAB33555	Aae33555 Human mic
42	739	22.4	688	7 ADG91726	Adg91726 Human mic
43	739	22.4	688	7 ADG91727	Adg91727 Human mic
44	739	22.4	688	8 ABB62131	Abb62131 Tumour-as
45	736	22.4	724	5 ABB04431	Abb04431 Murine ne

ALIGNMENTS

RESULT 1					
AA50056	AAB50056 standard; protein; 631 AA.				
XX	ID				
XX	AAB50056;				
XX	AC				
XX	19-MAR-2001 (first entry)				
XX	DE				
XX	Murine lymph node Stromal cell kinase 1.				
XX	KW				
XX	wound healing; periodontal disease; inflammatory disease; tumour;				
XX	infection; allergy.				
XX	OS				
XX	Mus musculue.				
XX	PN				
XX	W0200073468-A1.				
XX	PD				
XX	07-DEC-2000.				
XX	PF				
XX	26-MAY-2000; 2000MO-US014696.				
XX	PR				
XX	28-MAY-1999; 99US-0136781P.				
XX	PA				
XX	(IMMV) IMMUNEX CORP.				
XX	PI				
XX	Bird TA, Virca GD, Martin U, Anderson DM;				
XX	WPI; 2001-061546/07.				
XX	N-PSDB; AAC90433.				
XX	PT				
XX	Novel murine and human kinase nucleic acids useful for treating				
XX	PT				
XX	inflammations, infections, tumors, allergies, autoimmune diseases, and				
XX	for stimulating or suppressing immune responses.				
XX	PS				
XX	Claim 10; Page 94-96; 106pp; English.				
XX	CC				
XX	The present sequence is Murine lymph node Stromal cell kinase 1 (MLSK-1).				
XX	CC				
XX	This protein is useful for treating a variety of disorders listed in the				
XX	disclosure of the specification, including autoimmune disorders, allergic				
XX	reactions, myeloid or lymphoid cell deficiencies, wound healing, and				
XX	tissue repair and replacement, burns, incisions and ulcers, periodontal				
XX	disease, inflammatory diseases, tumours and bacterial, viral or fungal				
XX	infection				
XX	SQ				
XX	Sequence 631 AA;				
XX	Query Match				
XX	100.0%; Score 3293; DB 4; Length 631;				

Best Local Similarity 100.0%; Pred. No. 1e-280;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQAPSASALASBARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
DB 1 MESVALLQRPSPQAPSASALASBARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
QY 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHITAIH 120
DB 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHITAIH 120
QY 121 EVFENSSTIYIWEYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIYHRDL 180
DB 121 EVFENSSTIYIWEYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIYHRDL 180
QY 181 KLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCSPLVASPEIVNGKPYGPEVDSWSL 240
DB 181 KLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCSPLVASPEIVNGKPYGPEVDSWSL 240
QY 241 GVLVYTLVHGTMFPDQDHTLVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
DB 241 GVLVYTLVHGTMFPDQDHTLVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
QY 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
DB 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
QY 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
DB 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
QY 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSSGYSSPEPSESGELLD 480
DB 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSSGYSSPEPSESGELLD 480
QY 481 ASDVFPVSGDPVEQKSPQASGLLHKRGILKLNKGFSTALEGTPPTFGSLDQLASSHPA 540
DB 481 ASDVFPVSGDPVEQKSPQASGLLHKRGILKLNKGFSTALEGTPPTFGSLDQLASSHPA 540
QY 541 ARPSRPSGAVSEDSILSSSEFPQDLPERLPETPLRGCVSVNLRGLQEPSPSEGLKRWQ 600
DB 541 ARPSRPSGAVSEDSILSSSEFPQDLPERLPETPLRGCVSVNLRGLQEPSPSEGLKRWQ 600
QY 601 ESLGDSCEFLTDQCEVTAAYRQALGICSKLS 631
DB 601 ESLGDSCEFLTDQCEVTAAYRQALGICSKLS 631

RESULT 2
ABO84756
ID ABO84756 standard; protein; 652 AA.
XX
AC ABO84756;
XX
DT 18-NOV-2004 (first entry)
XX
DB Murine cancer-associated protein (CAP) M907-004.
XX
KW Mouse; cancer-associated protein; CAP; cancer; cytostatic.
XX
OS Mus musculus.
XX
PN M02004058146-A2.
XX
PD 15-JUL-2004.
XX
PF 15-DEC-2003; 2003MO-US040081.
XX
PR 17-DEC-2002; 2002US-00322281.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;

XX WI, 2004-499109/47.
DR N-PSDB; ABD33083.
XX
PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Disclosure, SEQ ID NO 23; 162pp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug. Involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CAP of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 652 AA;
Query Match 100.0%; Score 3293; DB 8; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.e-280;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQAPSASALASBARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 60
DB 22 MESVALLQRPSPQAPSASALASBARPLADGLIKSPKPLMKQAVKHHKHNLRHYEFL 81
QY 61 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHITAIH 120
DB 82 ETLGKGTGKVKKARSSGRVLAIKSRKDKIDEODLHIREIETIMSSLNPHITAIH 141
QY 121 EVFENSSTIYIWEYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIYHRDL 180
DB 142 EVFENSSTIYIWEYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIYHRDL 201
QY 181 KLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCSPLVASPEIVNGKPYGPEVDSWSL 240
DB 202 KLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCSPLVASPEIVNGKPYGPEVDSWSL 261
QY 241 GVLVYTLVHGTMFPDQDHTLVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 300
DB 262 GVLVYTLVHGTMFPDQDHTLVKQISNGAYREPPKSDACGLIRWLLMVPTRRATLED 321
QY 301 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 360
DB 322 VASHMVMWNGYTTGVGEQELRSGHPSGDFGRASMDMLRRSSRPLLENGAKVCSFFKQ 381
QY 361 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 420
DB 382 HYPGGSTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPKSSLLPKGILKKSS 441
QY 421 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSSGYSSPEPSESGELLD 480
DB 442 TSSGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSSGYSSPEPSESGELLD 501
QY 481 ASDVFPVSGDPVEQKSPQASGLLHKRGILKLNKGFSTALEGTPPTFGSLDQLASSHPA 540
DB 502 ASDVFPVSGDPVEQKSPQASGLLHKRGILKLNKGFSTALEGTPPTFGSLDQLASSHPA 561
QY 541 ARPSRPSGAVSEDSILSSSEFPQDLPERLPETPLRGCVSVNLRGLQEPSPSEGLKRWQ 600
DB 562 ARPSRPSGAVSEDSILSSSEFPQDLPERLPETPLRGCVSVNLRGLQEPSPSEGLKRWQ 621
QY 601 ESLGDSCEFLTDQCEVTAAYRQALGICSKLS 631

Db 622 ESIGDSFSLTDCQEVTAAYRQALGICSKLS 652

RESULT 3

AAE19885

AAE19885 standard; protein; 630 AA.

18-JUN-2002 (first entry)

Rat SNF1/AMPK-Related Kinase (SNARK) protein.

Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes; glucose deprivation; lipid metabolism; therapy; lipoprotein disorder; hyperglycaemia; drug screening; hypoglycaemia.

Rattus sp.

Location/Qualifiers

Key Binding-site /note="Protein kinase ATP-binding region signature"

Domain /note="Serine/threonine kinase catalytic domain"

Active-site /note="Serine/threonine protein kinase active-site"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

Domain /note="Serine/threonine kinase catalytic domain"

CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and

CC in other cell types such as heart and skeletal muscles, as well as

CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted

CC to have insulin-like effects that would enhance the disposal of glucose

CC into muscle and reduce plasma glucose for the treatment of diabetes and

CC some type of disorders of lipoprotein production leading to increased

CC levels of cholesterol or triglycerides. SNARK or its variants may be

CC administered to a subject to treat or prevent a disease associated with

CC decreased expression of SNARK, such as diabetes. SNARK antibodies are

CC used to modulate SNARK activity either in vivo for therapeutic purposes,

CC or in vitro, for drug screening and related investigational purposes.

CC SNARK antagonists may be administered to increase fuel production,

CC decrease glucose uptake and increase levels of blood glucose in a patient

CC suffering from hypoglycaemia. The present sequence is rat SNARK protein

CC

Sequence 630 AA;

Query Match 93.2%; Score 3068.5; DB 5; Length 630;

Best Local Similarity 94.0%; Pred. No. 6; 2e-261;

Matches 593; Conservative 12; Mismatches 25; Indels 1; Gaps 1;

1 MESVALORPSQASASALASASARPLADGLIKSPKMKQAVKHHKNNRRHREFL 60

1 MESVALHRRKONLAPASALATESARPLADRLIKSPKMKQAVKHHKNNRRHREFL 60

61 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHITAH 120

61 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHITAH 120

121 EVFENSKIIVMEVARSGLDYISERPLSRDARHFRFQIVSALHYCHONGIYARDL 180

121 EVFENSKIIVMEVARSGLDYISERPLSRDARHFRFQIVSALHYCHONGIYARDL 180

181 KLENTLLDNGNKIKIDFGISNLYHKGKPLQTCGSPIVASPEIVNGKPVVGEVNSWL 240

181 KLENTLLDNGNKIKIDFGISNLYHKGKPLQTCGSPIVASPEIVNGKPVVGEVNSWL 240

241 GVLLYILVHGTMPFDGQDHKTIVKQISNGAYRPPSPDACGLIRWLLMNPTRATLED 300

241 GVLLYILVHGTMPFDGQDHKTIVKQISNGAYRPPSPDACGLIRWLLMNPTRATLED 300

301 VASHMVMNMGYTTGVEBQELRGGHPSGDPGRASNAWDIIRSSRPILLENGAYCSFFKQ 360

301 VASHMVMNMGYTTGVEBQELRGGHPSGDPGRASNAWDIIRSSRPILLENGAYCSFFKQ 360

361 HVPGGGSGTVPGLERQHSIKSRKENDMAQTLQNDPVEDTSSRGKNSLKPGLIKKKAS 420

361 HVPGGGSGTVPGLERQHSIKSRKENDMAQTLQNDPVEDTSSRGKNSLKPGLIKKKAS 420

421 TSSEGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSEGYSSSPSESGELTD 480

421 TSSEGEVQEDPQELRPVDPFGQVPVAVSLPRKGIILKKSQRSEGYSSSPSESGELTD 480

481 ASDVPSGDPVEQKSPQASGLLHRRGIILKNGKFSRTALEGTPSTFGSLDQASHPA 540

481 ASDVPSGDPVEQKSPQASGLLHRRGIILKNGKFSRTALEGTPSTFGSLDQASHPA 540

541 ARPSRPSGAVSESIISSESPDQDLPERRLPETPLAGCYSVNDLRGLRPPSGGLKRWQ 600

541 ARPSRPSGAVSESIISSESPDQDLPERRLPETPLAGCYSVNDLRGLRPPSGGLKRWQ 600

601 ESIGDSFSLTDCQEVTAAYRQALGICSKLS 631

601 ESIGDSFSLTDCQEVTAAYRQALGICSKLS 631

600 ESIGDSRFSILTDCQEVTAAYRQALGICSKLS 630

600 ESIGDSRFSILTDCQEVTAAYRQALGICSKLS 630

RESULT 4

AAAM93360

AAAM93360 standard; protein; 628 AA.

AAAM93360;

06-NOV-2001 (first entry)

DB 474 AGDVAVSGDPEKQKPPQASGLLHRRKGLTLNKGSTALFLAAPTFTGSLDELAPRRPL 533
QY 541 ARPSRPSGAVSEDSIISSSEFPQDLPERLPETPLRGCVSVNLRGLRPPSG---LK 596
DB 534 ARASRPSGAVSEDSIISSSEFPQDLPERLPETPLRGCVSVNLRGLRPPSGSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 594 RWRQDPLGDSGCSFLTDCQEVTAAYRQALRVCSKLT 628

RESULT 7
AAU79652
AAU79652 standard; protein; 628 AA.
XX AAU79652;
DT 02-JUL-2002 (first entry)
XX
DE Human protein kinase 3700.
XX
KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
KW cell proliferation disorder; cell differentiation disorder; carcinoma;
KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;
KW cytoskeletal; antiatherosclerotic; enzyme.
XX
OS Homo sapiens.
XX
PN MO20024921-A2.
XX
PD 28-MAR-2002.
XX
PF 25-SEP-2001; 2001MO-US030115.
XX
PR 25-SEP-2000; 2000US-0234922P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Galvin KM;
XX
DR WPI; 2002-352007/38.
XX
DR N-PSDB; ABK14000.
XX
PT Use of modulators of activity of 3700 protein for making medicament for
PT e.g., modulating protein phosphorylation or cell signaling, or for
PT treating or preventing cellular proliferative and/or differentiative
PT disorders.
XX
PS Claim 19; Fig 1; 115pp; English.
XX
XX The present invention relates to the isolation of a novel human protein
XX kinase designated 3700, and the polynucleotide sequence encoding it. The
XX invention also describes the use of a modulator of the activity of
XX protein kinase (PK) 3700 for making a medicament or pharmaceutical
XX composition for modulating the ability of a cell to phosphorylate an
XX amino acid residue of a substrate protein. Modulators of protein kinase
XX 3700 activity are useful for modulating protein phosphorylation, cell
XX signalling, tumorigenesis, mitogenesis, transcription of a gene,
XX angiogenesis, tissue repair, tissue regeneration, establishment or
XX progression of atherosclerosis, and signalling across the blood-brain
XX barrier. The polynucleotide and polypeptide molecules for protein kinase
XX 3700 may be used as diagnostic targets and therapeutic agents for
XX prognosticating, diagnosing, preventing, inhibiting, alleviating, or
XX curing PK-related disorders and cellular proliferative and/or
XX differentiative disorders (e.g. haematopoietic neoplastic disorders,
XX carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
XX polynucleotide sequence can be used to express protein kinase 3700, to
XX detect a genetic alteration in a 3700 gene, in chromosome mapping, for
XX tissue typing, in forensic biology, and as surrogate markers. The present
XX sequence represents human protein kinase 3700

XX
SQ Sequence 628 AA;
Query Match 85.8%; Score 2824.5; DB 5; Length 628;
Best Local Similarity 85.7%; Pred. No. 2e-239;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLQRPQAPSAALASASARPLADGLIISPPPLMKKQAVKHHHKNRHHRYEFL 60
DB 1 MESLVAPRSGPTPS---AAILARPLAAGLTKSPKPKKKQAVKHHHKNRHHRYEFL 56
QY 61 ETGKGTGKVKKARSSGRLVAIKSIRKDKIADQEDLLHIREIRIEMSLNPHIAT 120
DB 57 ETGKGTGKVKKARSSGRLVAIKSIRKDKIADQEDLLHIREIRIEMSLNPHIAT 116
QY 121 EVFENSSKIIVIMEVYASRGDLYDIISERPLSERDARHFRQIVSAHYCHONGIYHRL 180
DB 117 EVFENSSKIIVIMEVYASRGDLYDIISERQOOLSEREARHFRQIVSAHYCHONGIYHRL 176
QY 181 KLENILLDANGNIKIDFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKTYGPEVDSWSL 240
DB 177 KLENILLDANGNIKIDFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKTYGPEVDSWSL 236
QY 241 GYLLYTLVHGTMPFDGQDHKTLYKQISNGAYREPPKPSDACGILWLVNPTFRATLTD 300
DB 237 GYLLYTLVHGTMPFDGHDHKLIVKQISNGAYREPPKPSDACGILWLVNPTFRATLTD 296
QY 301 VASHMWNVNGYTTGVGEQALREGGHPSGDFGRASNADWLRRSSRPLTENGAVCSPFYQ 360
DB 297 VASHMWNVNGYATRVGEQALREGGHPSGDSARASNADWLRRSSRPLTENGAVCSPFYQ 356
QY 361 HYPGGGSTVPGIERQSLTKSRKENDMAQNLQDPAEDTSRPPGKSLTLPGIKLKKKS 420
DB 357 HAPGGGSTVPGIERQSLTKSRKENDMAQSLDSTDAHPRGKSNLTLPGIKLKKKS 416
QY 421 TSSGEVQEDPOELRPVPTFRGQVPVPAVSLPRKGIKKRSORESGYSSPPSESGELTD 480
DB 417 ASHGVQEDPEELSPVPASPGQAP--LPPKGIKKRQREBSGYSSPPSESGELTD 473
QY 481 ASDVAVSGDPEVOKSPQASGLLHRRKGLTLNKGSTALFLAAPTFTGSLDELASHPA 540
DB 474 AGDVAVSGDPEKQKPPQASGLLHRRKGLTLNKGSTALFLAAPTFTGSLDELAPRRPL 533
QY 541 ARPSRPSGAVSEDSIISSSEFPQDLPERLPETPLRGCVSVNLRGLRPPSG---LK 596
DB 534 ARASRPSGAVSEDSIISSSEFPQDLPERLPETPLRGCVSVNLRGLRPPSGSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
DB 594 RWRQDPLGDSGCSFLTDCQEVTAAYRQALRVCSKLT 628

RESULT 8
ADF76965
ADF76965 standard; protein; 628 AA.
XX ADF76965;
AC ADF76965;
XX
DT 26-FEB-2004 (first entry)
XX
DE Novel human secreted and transmembrane protein SeqID 640.
XX
XX human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytoskeletal; chondrocyte; tumour.
XX
OS Homo sapiens.
XX
XX
XX
PN MO2003072035-A2.
XX
PD 04-SEP-2003.

Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;			
QY	1	MESVALLQRPSPQPSAALASASARPADGLIKSPKLMKQAVKRRHHKNIHRHYEFL	60
DB	1	MESLVARRSGPTPS---AAELARPLAELGLIKSPKLMKQAVKRRHHKNIHRHYEFL	56
QY	61	ETLGKTYGVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH	120
DB	57	ETLGKTYGVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH	116
QY	121	EYFENSSKIYIVMEYASRGDLVDYISERPLSERDARHFRQIVSAVHYCHQRRVYHRL	180
DB	117	EYFENSSKIYIVMEYASRGDLVDYISERPLSERDARHFRQIVSAVHYCHQRRVYHRL	176
QY	181	KENITLLDANGINKIADFGLSNIYHKGKFLQTFCCGSLVYASPEIVNGKPYGPEVDSWSL	240
DB	177	KENITLLDANGINKIADFGLSNIYHKGKFLQTFCCGSLVYASPEIVNGKPYGPEVDSWSL	236
QY	241	GVLLYTLVHGTMFPDGDHDKTLVKQISNGAYRBPSPDAGLIRWLLMVPTRRATLTD	300
DB	237	GVLLYTLVHGTMFPDGDHDKTLVKQISNGAYRBPSPDAGLIRWLLMVPTRRATLTD	296
QY	301	VASHMWMNGYTTGVGEQALREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ	360
DB	297	VASHMWMNGYATRVGEQALREGGHPSGDSARASMAWMLRRSSRPLLENGAKVCSFFKQ	356
QY	361	HVPGGGSLVPGLEKRSILKSKRENDMAQNLQDPAEDTSSRPGKSLKPKGILLKKKS	420
DB	357	HAPGGGSLVPGLEKRSILKSKRENDMAQNLQDPAEDTSSRPGKSLKPKGILLKKKS	416
QY	421	TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRRQSGYSSPEPSESEELD	480
DB	417	ASAEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRRQSGYSSPEPSESEELD	473
QY	481	ASDVFSVGPVPEKSPQASGLLHRRKGIILKNGKFSRTALEGTPPTFGSLDOLASHPA	540
DB	474	AGDVFSVGPVPEKSPQASGLLHRRKGIILKNGKFSRTALEGTPPTFGSLDOLASHPA	533
QY	541	ARSRPSGAVSEDSILSSSEFDOLDLPERLPETPLRGCVSVNLRGLQEPSESG---	596
DB	534	ARSRPSGAVSEDSILSSSEFDOLDLPERLPETPLRGCVSVNLRGLQEPSESG---	593
QY	597	RMWQSLGDSGSLTDCQEVTAAYROALGICSKLS	631
DB	594	RMWQSLGDSGSLTDCQEVTAAYROALGICSKLT	628
RESULT 10			
ADL25362			
ADL25362 standard; protein; 628 AA.			
AC ADL25362;			
DT 03-JUN-2004 (first entry)			
DE Human SNARK, SEQ ID 12.			
KW Cytostatic; Neuroprotective; Nootropic; Muscular; Gene therapy;			
KW Adenosine Monophosphate activated Protein Kinase-related kinase 5;			
KW AMPK-related kinase 5; ARKS; KIAA0537; stress resistance; tumor;			
KW nervous disorder; muscle disorder; ataxia telangiectasia; SNARK.			
OS Homo sapiens.			
PN WO2004019994-A1.			
PD 11-MAR-2004.			
PF 19-AUG-2003; 2003WO-IP010435.			
PR 27-AUG-2002; 2002JP-00247761.			
PA (NINA-) JAPAN AGENCY NAT INST HEALTH.			

PA	(PHAR-) ORG PHARM SAFETY & RES.
XX	PI Esumi H, Suzuki A;
XX	DX WPI; 2004-248195/23.
XX	PT Agent for imparting stress resistance to cells, comprises DNA encoding
PT	for adenosine monophosphate activated protein kinase-related kinase 5 for
PT	treatment of tumors, muscle and nervous disorder, and ataxia
PS	teleangiectasia.
XX	Example 1; SEQ ID NO 12; 143pp; Japanese.
XX	CC The present invention relates to human Adenosine Monophosphate activated
CC	Protein Kinase (AMPK)-related kinase 5 (ARKS/ KIA0537) and its coding
CC	sequence, which are useful for imparting stress resistance to cells. ARKS
CC	is useful for treatment and prevention of tumors, nervous disorders,
CC	muscle disorders and ataxia teleangiectasia. The present sequence was
CC	used to illustrate the invention.
XX	Sequence 628 AA;
SQ	
Query Match	85.8%; Score 2824.5; DB 8; Length 628;
Best Local Similarity	85.7%; Pred. No. 2e-239;
Matches 544; Conservative	26; Mismatches 54; Indels 11; Gaps 3;
QY	1 MESVALLQRPSPQPSAALASASARPADGLIKSPKLMKQAVKRRHHKNIHRHYEFL 60
DB	1 MESLVARRSGPTPS---AAELARPLAELGLIKSPKLMKQAVKRRHHKNIHRHYEFL 56
QY	61 ETLGKTYGVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 120
DB	57 ETLGKTYGVKKARSSGRLVAIKSIRKDKIODEDLHIREIEMSLNPHITAIH 116
QY	121 EYFENSSKIYIVMEYASRGDLVDYISERPLSERDARHFRQIVSAVHYCHQRRVYHRL 180
DB	117 EYFENSSKIYIVMEYASRGDLVDYISERPLSERDARHFRQIVSAVHYCHQRRVYHRL 176
QY	181 KLENITLLDANGINKIADFGLSNIYHKGKFLQTFCCGSLVYASPEIVNGKPYGPEVDSWSL 240
DB	177 KLENITLLDANGINKIADFGLSNIYHKGKFLQTFCCGSLVYASPEIVNGKPYGPEVDSWSL 236
QY	241 GVLLYTLVHGTMFPDGDHDKTLVKQISNGAYRBPSPDAGLIRWLLMVPTRRATLTD 300
DB	237 GVLLYTLVHGTMFPDGDHDKTLVKQISNGAYRBPSPDAGLIRWLLMVPTRRATLTD 296
QY	301 VASHMWMNGYTTGVGEQALREGGHPSGDFGRASMAWMLRRSSRPLLENGAKVCSFFKQ 360
DB	297 VASHMWMNGYATRVGEQALREGGHPSGDSARASMAWMLRRSSRPLLENGAKVCSFFKQ 356
QY	361 HVPGGGSLVPGLEKRSILKSKRENDMAQNLQDPAEDTSSRPGKSLKPKGILLKKKS 420
DB	357 HAPGGGSLVPGLEKRSILKSKRENDMAQNLQDPAEDTSSRPGKSLKPKGILLKKKS 416
QY	421 TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRRQSGYSSPEPSESEELD 480
DB	417 ASAEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRRQSGYSSPEPSESEELD 473
QY	481 ASDVFSVGPVPEKSPQASGLLHRRKGIILKNGKFSRTALEGTPPTFGSLDOLASHPA 540
DB	474 AGDVFSVGPVPEKSPQASGLLHRRKGIILKNGKFSRTALEGTPPTFGSLDOLASHPA 533
QY	541 ARSRPSGAVSEDSILSSSEFDOLDLPERLPETPLRGCVSVNLRGLQEPSESG---LK 596
DB	534 ARSRPSGAVSEDSILSSSEFDOLDLPERLPETPLRGCVSVNLRGLQEPSESG---CLR 593
QY	597 RMWQSLGDSGSLTDCQEVTAAYROALGICSKLS 631
DB	594 RMWQSLGDSGSLTDCQEVTAAYROALGICSKLT 628
RESULT 11	
ADL14161	

ID ADL14161 standard; protein; 628 AA.
 AC ADL14161;
 XX
 XX
 DT 17-JUN-2004 (first entry)
 DE Novel human gene 3700 encoded protein.
 XX
 XX cytostatic; cardiac; hypotensive; antianginal; osteopathic;
 KM antithrombotic; antirheumatic; neuroprotective; antiinflammatory;
 KM angioplastic; antihemorrhagic; cardiovascular; vitruclide; analgesic; CNS;
 KM angiotensin inhibitor; angiotensin stimulator; cerebroprotective;
 KM nephrotoxic; antithyroid; dermatological; immunomodulator;
 KM cell proliferation disorder; cell differentiation disorder;
 KM brain disorder; platelet disorder; breast disorder; colon disorder;
 KM kidney disorder; renal disorder; lung disorder; ovarian disorder;
 KM prostate disorder; cervical disorder; spleen disorder; thymus disorder;
 KM thyroid disorder; testes disorder; hematopoietic disorder;
 KM pancreatic disorder; bone metabolism disorder; skin disorder;
 KM dermal disorder; bone metabolism disorder; immune disorder;
 KM inflammatory disorder; cardiovascular disorder; viral disease; pain disorder;
 KM endothelial cell disorder; liver disorder;
 KM metabolic disorder; neurological disorder;
 KM central nervous system disorder; erythroid disorder;
 KM blood vessel disorder; angiogenic disorder; cancer; heart failure;
 KM hypertension; angina; osteoarthritis; rheumatoid arthritis;
 KM multiple sclerosis; Crohn's disease; psoriasis; asthma;
 KM cell proliferation; cell differentiation; cell growth; cell division;
 KM human.
 XX
 XX Homo sapiens.
 OS
 XX
 XX US2004058355-A1.
 PD
 XX 25-MAR-2004.
 XX
 XX 25-APR-2003; 2003US-00423543.
 XX
 XX 30-SEP-1998; 98US-00163821.
 PR 27-JUN-1999; 99US-0117580P.
 PR 25-MAR-1999; 99US-00276400.
 PR 30-JUL-1999; 99US-00365162.
 PR 09-SEP-1999; 99US-00392189.
 PR 05-OCT-1999; 99US-00412210.
 PR 23-NOV-1999; 99US-00448076.
 PR 29-FEB-2000; 2000US-0186061P.
 PR 28-APR-2000; 2000US-0206688P.
 PR 19-MAY-2000; 2000US-0205447P.
 PR 30-JUN-2000; 2000US-00608921.
 PR 31-JUL-2000; 2000US-0221925P.
 PR 25-SEP-2000; 2000US-0234922P.
 PR 25-SEP-2000; 2000US-0235035P.
 PR 08-NOV-2000; 2000US-024669P.
 PR 09-NOV-2000; 2000US-00711216.
 PR 14-NOV-2000; 2000US-0248325P.
 PR 15-NOV-2000; 2000US-0248893P.
 PR 22-DEC-2000; 2000US-0257511P.
 PR 05-JAN-2001; 2001US-0260166P.
 PR 28-FEB-2001; 2001US-00797039.
 PR 27-APR-2001; 2001US-00845044.
 PR 20-JUL-2001; 2001US-00909743.
 PR 31-JUL-2001; 2001US-00920346.
 PR 13-AUG-2001; 2001US-00928531.
 PR 14-AUG-2001; 2001US-00929218.
 PR 15-AUG-2001; 2001US-0312539P.
 PR 25-SEP-2001; 2001US-00963159.
 PR 08-NOV-2001; 2001US-00008016.
 PR 13-NOV-2001; 2001US-00012055.
 PR 15-NOV-2001; 2001US-00003690.
 PR 30-JAN-2002; 2002US-00060763.
 PR 25-MAR-2002; 2002US-00105989.
 PR 12-APR-2002; 2002US-00121911.
 PR 12-AUG-2002; 2002US-00217168.

PR 22-OCT-2002; 2002US-00278036.
 PR 02-JAN-2003; 2003US-00336489.
 PR 03-JAN-2003; 2003US-00336153.
 XX
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Kapeller-Libermann R, Hunter JI, Meyers RE, Rudolph-Owen LA;
 PI Curtis RWJ, Olandt PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;
 PI Siles-Santiego I, Bandaru R;
 XX
 DR WPI: 2004-268788/25.
 DR N-PSDB; ADL14160, ADL14162.
 XX
 XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,
 PT 26176, 26343, 56638, 18610, 33217, 21967, 11983, 38555 or 593
 PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,
 PT heart failure and angina.
 XX
 PS Claim 4; SEQ ID NO 44; 139pp; English.
 XX
 XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,
 CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
 CC 21967, 11983, 38555 or 593 nucleic acid molecule (I) comprising
 CC any one of 40 nucleotide sequences (I). The nucleic acid molecules and
 CC polypeptides are useful for diagnosing and treating a subject having a
 CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
 CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, 11983,
 CC m1983, 38555 or 593 activity, such as cellular proliferative and/or
 CC differentiative disorders, brain disorders, platelet disorders, breast
 CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
 CC ovarian disorders, prostate disorders, cervical disorders, spleen
 CC disorders, thymus disorders, thyroid disorders, testes disorders,
 CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
 CC skin (dermal) disorders, disorders associated with bone metabolism,
 CC immune, e.g. inflammatory disorders, cardiovascular disorders, pain
 CC disorders, endoheial cell disorders, liver disorders, viral diseases, pain
 CC disorders, metabolic disorders, neurological or central nervous system
 CC disorders, erythroid disorders, blood vessel disorders or angiogenic
 CC disorders (all claimed), e.g. cancer, heart failure, hypertension, Crohn's
 CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
 CC disease, psoriasis, or asthma. The nucleic acid molecules and
 CC polypeptides are also useful as modulating agents in regulating a variety
 CC of cellular processes, e.g. cell proliferation, differentiation, growth and
 CC division. This is the amino acid sequence of a novel human protein of the
 CC invention. Note: The sequences given in the specification are also
 CC available in electronic format from
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20040058355.
 CC
 XX
 SQ Sequence 628 AA;
 Query Match 85.8%; Score 2824.5; DB 8; Length 628;
 Best Local Similarity 85.7%; Pred. No. 2e-23;
 Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
 QY 1 MESVALORPSQASALASASARPLADGLISPRIMKKQAVKHHKNNRRHREFL 60
 DB 1 MESLVFARRSGPTPS---AAELARPLAELISPRPLMKQAVKHHKNNRRHREFL 56
 QY 61 ETLGKGTGKVKKARRSGRLVAIKSIRKDKIODEBDLHIREIEMSLNPHIIAH 120
 DB 57 ETLGKGTGKVKKARRSGRLVAIKSIRKDKIODEBDLHIREIEMSLNPHIIAH 116
 QY 121 EYFENSSKIYIVMEYASRGDLVYISERPLSRDRAHPRFQVSAHYHONGIYARDL 180
 DB 117 EYFENSSKIYIVMEYASRGDLVYISERPLSRDRAHPRFQVSAHYHONGIYARDL 176
 QY 181 KLENIIIDANGNKIKIDFGSLNLYHKGFQTCGSPLYASPRITNGKPYGPEVNSWSL 240
 DB 177 KLENIIIDANGNKIKIDFGSLNLYHKGFQTCGSPLYASPRITNGKPYGPEVNSWSL 236
 QY 241 GVLVYLIVHGTMPDGDHRTLVKQISNGAYRPPKPSDAGLIRMLMNVNTPRATLED 300

Db 237 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKSDACGLIRMLMWNPTRRATLED 296
QY 301 VASHMWMVNGYTTGVBGEQALREGCHPGSDGFSASADWLRRSSRPILLENGAKVCSFFKQ 360
Db 297 VASHMWMVNGYATRVGEQAPHEGCHPGSDSARASADWLRRSSRPILLENGAKVCSFFKQ 356
QY 361 HVGGSSTVYGLERHSLKRSKRENDMAQNLGDPADPTSSRPCKSSLKLPKGIILKKKS 420
Db 357 HAPGGSTTPGLERHSLKRSKRENDMAQNLGDPADPTSSRPCKSSLKLPKGIILKKKS 416
QY 421 TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRQSRGYSSPEPSESGELLD 480
Db 417 ASAEVQEDPPELSPIPASPGQAP---LLPKGILKKPRQSRGYSSPEPSESGELLD 473
QY 481 ASDVFGSDPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPPTFGSLDQLASHPA 540
Db 474 AGDVFGSDPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPPTFGSLDQLASHPA 533
QY 541 ARPSRPSGAVSEDSILSSSEFPDQLDPERLPETPLRGCVSVNLRGLBOPPSG---LK 596
Db 534 ARPSRPSGAVSEDSILSSSEFPDQLDPERLPETPLRGCVSVNLRGLBOPPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 12

ADO20172
ID ADO20172 standard; protein; 628 AA.

AC ADO20172;

DT 12-AUG-2004 (first entry)

DE Human PRO polypeptide #540.

Human; PRO; immune related disorder; systemic lupus erythematosus;
rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
diabetes mellitus; renal disease; demyelinating disease;
central nervous system; peripheral nervous system;
demyelinating polyneuropathy; Guillain-Barre syndrome;
chronic inflammatory demyelinating polyneuropathy.

OS Homo sapiens.

PN MO2004043361-A2.

PD 27-MAY-2004.

PF 06-NOV-2003; 2003WO-US035268.

PR 08-NOV-2002; 2002US-0425235P.

PA (GERTH) GENENTECH INC.

PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PW,
PI Wood WI, Wu TD;

DR MPI; 2004-420067/39.

DR N-PSDB; ADO20171.

Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO68388 useful for
treating an immune related disorder such as systemic lupus erythematosus,
rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
spondyloarthritis.

Claim 7; SEQ ID NO 1080; 1731pp; English.

The invention relates to human PRO polypeptides and the polynucleotides
encoding them. The polypeptides and polynucleotides are useful for

treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.

SQ Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;

Best Local Similarity 85.7%; Pred. No. 26-239;

Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLORPSQAPSALASASAPRLADGLIKSPKPLMKQAVRHHKHLRHHYEPL 60
Db 1 MESLVFARRSGTPPS---AAELARPLAEGLIKSPKPLMKQAVRHHKHLRHHYEPL 56
QY 61 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEODLHIREIEINSLNHPHIIAT 120
Db 57 ETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEODLHIREIEINSLNHPHIIAT 116
QY 121 EVFENSSKIYIWEIASRGDLYDIYISERPLSERDARHPFROI VSAIHYCHONGIVHRDL 180
Db 117 EVFENSSKIYIWEIASRGDLYDIYISERQULSEREARHFRQIVSAVHYCHONRVHRDL 176
QY 181 KLENILLDANGNIKIADPGLSNLYHKHGFLOFCGSLPYASPEINVGKRYVPEVDSWLT 240
Db 177 KLENILLDANGNIKIADPGLSNLYHOGKFLQTFCCSPLASPEINVGKRYTPEVDSWLT 236
QY 241 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKSDACGLIRMLMWNPTRRATLED 300
Db 237 GVLLYLIVHGTMPFDGHDHKLIVKQISNGAYRBPCKSDACGLIRMLMWNPTRRATLED 296
QY 301 VASHMWMVNGYTTGVBGEQALREGCHPGSDGFSASADWLRRSSRPILLENGAKVCSFFKQ 360
Db 297 VASHMWMVNGYATRVGEQAPHEGCHPGSDSARASADWLRRSSRPILLENGAKVCSFFKQ 356
QY 361 HVGGSSTVYGLERHSLKRSKRENDMAQNLGDPADPTSSRPCKSSLKLPKGIILKKKS 420
Db 357 HAPGGSTTPGLERHSLKRSKRENDMAQNLGDPADPTSSRPCKSSLKLPKGIILKKKS 416
QY 421 TSSGEVQEDPQELRPVDPTRGQVPAVSLPRKGIILKKSRQSRGYSSPEPSESGELLD 480
Db 417 ASAEVQEDPPELSPIPASPGQAP---LLPKGILKKPRQSRGYSSPEPSESGELLD 473
QY 481 ASDVFGSDPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPPTFGSLDQLASHPA 540
Db 474 AGDVFGSDPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPPTFGSLDQLASHPA 533
QY 541 ARPSRPSGAVSEDSILSSSEFPDQLDPERLPETPLRGCVSVNLRGLBOPPSG---LK 596
Db 534 ARPSRPSGAVSEDSILSSSEFPDQLDPERLPETPLRGCVSVNLRGLBOPPSGCLR 593
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 13

ABO84757
ID ABO84757 standard; protein; 628 AA.

AC ABO84757;

DT 18-NOV-2004 (first entry)

DE Human cancer-associated protein (CAP) HP07-004.

Human; cancer-associated protein; CAP; cancer; cytosstatic.

XX Homo sapiens.
 XX MO2004058146-A2.
 XX
 PD 15-JUL-2004.
 XX
 PF 15-DEC-2003; 2003MO-US040081.
 XX
 PR 17-DEC-2002; 2002US-00322281.
 XX
 PA (SAGR-) SAGRES DISCOVERY INC.
 PI Morris DW, Malandro MS;
 DR WPI; 2004-499109/47.
 DR N-PSDB; ABDJ3085.
 XX
 PT Novel human cancer associated protein encoded within open reading frame
 of cancer associated gene, useful as targets for diagnosing cancer.
 XX
 PS Claim 18; SEQ ID NO 26; 182bp; English.
 XX
 CC The invention relates to cancer-associated proteins (CAP) and the cancer-
 associated (CA) nucleic acids encoding them. The invention also relates
 to a method for treating cancers involving administering to a patient an
 inhibitor of CAP, and a method of screening for anticancer activity in a
 potential drug involving providing a cell that expresses a CA gene,
 contacting a tissue sample derived from a cancer cell with an anticancer
 drug candidate and monitoring the effect of the anticancer drug candidate
 on expression of the CA gene. The CAP proteins are useful for detecting
 cancer associated with expression of a CAP protein in a test cell sample
 and for screening for a bioactive agent capable of modulating the
 activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 cancer, involving determining the expression of a CA nucleic acid in a
 tissue. This sequence represents a human CAP of the invention. Note: The
 sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 628 AA:
 Query Match 85.8%; Score 2824.5; DB 8; Length 628;
 Best Local Similarity 85.7%; Pred. No. 26-239;
 Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
 QY 1 MESVALLQRPSPQASALASASAPRLADGLISPPRLMKQAQVKKHHKHNIRHYEFL 60
 DB 1 MESLVFARRSGPTPS---AAELARPLAEGLIKSPKPLMKQAQVKKHHKHNIRHYEFL 56
 QY 61 ETLGKGTGKVKKARRSGRLVAIKIRKDKIDBDLHIREIEMSLNPHITAH 120
 DB 57 ETLGKGTGKVKKARRSGRLVAIKIRKDKIDBDLHIREIEMSLNPHITAH 116
 QY 121 EYFENSSTIVIVMEVARSGLDYISERPRLSRDARHFRQIVASALHYCHQGIYHRDL 180
 DB 117 EYFENSSTIVIVMEVARSGLDYISERQLSREARHFRQIVASAVHYCHQRIYHRDL 176
 QY 181 KLENIILLDANGNITKIDPGLISNLVHKGKLTQFCGSPLYASPIVNGKPYGPEVDSWSL 240
 DB 177 KLENIILLDANGNITKIDPGLISNLVHKGKLTQFCGSPLYASPIVNGKPYGPEVDSWSL 236
 QY 241 GVLVLYLVGTMFPDGDHKTIVKOISNGAYRPPPSDSCGILRWLMVNPFRRTATLEB 300
 DB 237 GVLVLYLVGTMFPDGDHKTIVKOISNGAYRPPPSDSCGILRWLMVNPFRRTATLEB 296
 QY 301 VASHMWNVNGYITTVGEQEARLREGHPSGDFGRASMAWMLRRSSRPLLENGATVCSFFPKQ 360
 DB 297 VASHMWNVNGYITTVGEQEARLREGHPSGDFGRASMAWMLRRSSRPLLENGATVCSFFPKQ 356
 QY 361 HYFGGSGTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPFGSKSLKPKGILLKKSS 420
 DB 357 HAFGGSGTVPGLERQHSLLKSRKENDMAQNLQDPAEDTSSRPFGSKSLKPKGILLKKSS 416

QY 421 TSSGEVQEDPQELRPVDTPPQVPAVSLPKKGIILKSRQRESGYSSPPESSGELLD 480
 DB 417 ASAEVQEDPPELSPFPASPGQAP---LLPKGIIILKSRQRESGYSSPPESSGELLD 473
 QY 481 ASDVFVSGDPEVCKSRQASGLLHRRGIIILKLNKFSRTALEGTTPTFGSLDGLASHPA 540
 DB 474 AGDVFVSGDPEVCKSRQASGLLHRRGIIILKLNKFSRTALEGTTPTFGSLDGLAPPL 533
 QY 541 ARPSRPGAVSEDSIILSSSFDQDLPERLPEPLRGCVSVNLRGLGEPSESG---LK 596
 DB 534 ARASRPGAVSEDSIILSSSFDQDLPERLPEPLRGCVSVNLRGLGEPSESGCLR 593
 QY 597 RWMQESLGDSGSLTDCQEVTAAYRQALGICSKLS 631
 DB 594 RWMQDPLGDSGSLTDCQEVTAAYRQALRVCSKLT 628
 RESULT 14
 ADJ96620
 ID ADJ96620 standard; protein; 672 AA.
 XX
 AC ADJ96620;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human calcium/calmodulin-dependent protein kinase NuaK2 protein SeqID 77.
 XX
 KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
 KW PK; ST; gene therapy; cancer; immune-related disease;
 KW cardiovascular disease; brain; neuronal associated disease; metabolic;
 KW inflammatory disorder; cytosolic; neuroprotective; immunomodulator;
 KW antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;
 KW NuaK.
 XX
 OS Homo sapiens.
 OS
 OS 72.
 XX
 PN MO2004006838-A2.
 XX
 PD 22-JUN-2004.
 XX
 PF 15-JUL-2003; 2003MO-US021730.
 XX
 PR 15-JUL-2002; 2002US-0395632P.
 XX
 PA (SUGR-) SUGEN INC.
 PI Whyte D, Manning G, Caenepeel S;
 XX
 DR WPI; 2004-122753/12.
 DR N-PSDB; ADJ96554.
 XX
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for
 preparing a composition for treating diseases or disorders, e.g., cancer,
 or neurological, immunological or inflammatory disorders.
 PT
 PS Claim 1; SEQ ID NO 77; 366bp; English.
 XX
 CC This invention relates to a novel isolated, enriched or purified nucleic
 acid molecule that encodes a kinase polypeptide. Specifically, it relates
 to human tyrosine and serine/threonine protein kinases (PK's and STK's),
 as well as protein kinase-like enzymes. The present invention describes
 screening methods to identify agonists, antagonists and antibodies that
 can be used to modulate the activity or function of the mammalian kinase
 enzymes. As such, these compositions can be used for gene therapy
 purposes to treat diseases or disorders including cancer, immune-related
 diseases, cardiovascular disease, brain or neuronal associated disease,
 CC metabolic and inflammatory disorders. Accordingly, they exhibit
 CC cytostatic, neuroprotective, immunomodulator and antiinflammatory
 CC activities. This polypeptide sequence is a human kinase protein sequence
 of the invention.
 XX

Oy	480	DASDVYFVSGDPVYBOKSPQASGILLHRKGIKLNKFSRTALBCTTPTSGSLDOLASSHP	539
Db	474	DAGDVYFVSGDPKEQKPPQASGILLHRKGIKLNKFSQTALBAPTTGSLDELAPPRP	533
Oy	540	AARPSRPSGAVSBDSTLSSSEFPQDLPERLPERPLRGCVSVNLTGLBOPPEG----	595
Db	534	LARASRPSGAVSBDSTLSSSEFPQDLPERLPERPLRGCVSVNLTGLBPPSEGGPSCL	593
Oy	596	KRMWESLGDSCFSLTDCOEVTAAVROALGICSKLS	631
Db	594	RRWRQDPLGDSCTSLTDCOEVTATYRQALRVCSKLT	629

Search completed: May 11, 2005, 14:20:07
 Job time : 171.602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:24:11 ; Search time 129.912 Seconds
(without alignments)
1622.537 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293
Sequence: 1 MESVALLQRPSPQAPSASALA.....DCQETVAYRQALGICSKLS 631

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications AA:*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	14	US-10-355-975-11 Sequence 11, Appl
2	3186	96.8	640	16	US-10-322-281-23 Sequence 23, Appl
3	3068.5	93.2	630	16	US-10-343-514-41 Sequence 41, Appl
4	2824.5	85.8	628	9	US-09-963-159-2 Sequence 2, Appl
5	2824.5	85.8	628	15	US-10-423-543-44 Sequence 44, Appl
6	2824.5	85.8	628	16	US-10-618-941-77 Sequence 77, Appl
7	2720.5	82.6	594	15	US-10-311-034-12 Sequence 12, Appl
8	2712.5	82.4	616	16	US-10-322-281-26 Sequence 26, Appl
9	1781.5	54.1	406	16	US-10-363-829-404 Sequence 404, App
10	1781.5	54.1	406	16	US-10-363-829-504 Sequence 504, App
11	1699.5	51.6	661	9	US-09-780-949-2 Sequence 2, Appl
12	1699.5	51.6	661	9	US-09-780-949-6 Sequence 2, Appl
13	1699.5	51.6	661	14	US-10-354-358-82 Sequence 82, Appl

14	1325	40.2	251	16	US-10-343-514-103	Sequence 103, App
15	1301.5	39.5	530	9	US-09-836-392-20	Sequence 20, Appl
16	1296	39.4	251	16	US-10-343-514-50	Sequence 50, Appl
17	1276	38.7	251	16	US-10-343-514-101	Sequence 101, App
18	1111.5	33.8	252	16	US-10-343-514-102	Sequence 102, App
19	785.5	23.9	246	10	US-09-898-837A-29	Sequence 29, Appl
20	748	22.7	703	15	US-10-016-248-70	Sequence 70, Appl
21	748	22.7	752	9	US-09-835-081-2	Sequence 2, Appl
22	748	22.7	752	15	US-10-258-106-16	Sequence 16, Appl
23	748	22.7	752	15	US-10-276-645-7	Sequence 7, Appl
24	748	22.7	825	15	US-10-425-114-54516	Sequence 54516, A
25	744	22.6	639	15	US-10-016-248-71	Sequence 71, Appl
26	744	22.6	639	15	US-10-016-248-72	Sequence 72, Appl
27	744	22.6	688	15	US-10-276-645-8	Sequence 8, Appl
28	744	22.6	752	16	US-10-618-941-79	Sequence 79, Appl
29	743.5	22.6	769	15	US-10-363-616-403	Sequence 403, Appl
30	739	22.4	639	15	US-10-016-248-73	Sequence 73, Appl
31	739	22.4	688	14	US-10-161-565-28	Sequence 28, Appl
32	739	22.4	688	14	US-10-161-565-29	Sequence 29, Appl
33	736	22.4	724	15	US-10-276-645-5	Sequence 5, Appl
34	732.5	22.2	660	15	US-10-276-645-6	Sequence 6, Appl
35	703.5	21.4	729	14	US-10-142-356-11	Sequence 11, Appl
36	703.5	21.4	729	14	US-10-195-101-33	Sequence 33, Appl
37	703.5	21.4	729	14	US-10-161-565-26	Sequence 26, Appl
38	701	21.3	140	16	US-10-343-514-47	Sequence 47, Appl
39	701	21.3	744	9	US-09-835-081-4	Sequence 4, Appl
40	700.5	21.3	1038	15	US-10-424-599-274878	Sequence 274878, A
41	698.5	21.2	624	15	US-10-425-114-70100	Sequence 70100, A
42	698.5	21.2	744	9	US-09-919-585-3	Sequence 3, Appl
43	698	21.2	698	15	US-10-016-248-22	Sequence 22, Appl
44	696.5	21.2	508	15	US-10-016-248-24	Sequence 24, Appl
45	694.5	21.1	667	15	US-10-016-248-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-10-355-975-11
; Sequence 11, Application US/103555975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/10/355,975
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579,664B
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-355-975-11

Query Match 100.0%; Score 3293; DB 14; Length 631;
Best Local Similarity 100.0%; Pred. No. 1.1e-206;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESVALLQRPSPQAPSASALASARPLADGLIKSPPLMKQAVKRRHHRNLRHYEFL 60
DB 1 MESVALLQRPSPQAPSASALASARPLADGLIKSPPLMKQAVKRRHHRNLRHYEFL 60
QY 61 ETIGKTYGVKVRARSSGRIVAIKIRKDKIKDODLHIREIIMSLNPHIAT 120
DB 61 ETIGKTYGVKVRARSSGRIVAIKIRKDKIKDODLHIREIIMSLNPHIAT 120

```
QY 121 EVFENSKIYIVMEYASRGDLVDYISERPRLSERDARHPFROIYSALHYCHONGIYHRDL 180
| | | | |
DB 121 EVFENSKIYIVMEYASRGDLVDYISERPRLSERDARHPFROIYSALHYCHONGIYHRDL 180
| | | | |
QY 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPFLYASPEIYNKGPVYGVSEVDSWSL 240
| | | | |
DB 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPFLYASPEIYNKGPVYGVSEVDSWSL 240
| | | | |
QY 241 GULLYILVHGTMFPDGDHKTLYKQISNGAYREPPKSDACGLIRWLMLNPTRRATLED 300
| | | | |
DB 241 GULLYILVHGTMFPDGDHKTLYKQISNGAYREPPKSDACGLIRWLMLNPTRRATLED 300
| | | | |
QY 301 VASHMMVNMNGYTTGVGEQELRGGHPSGDFGRASMDWLRRSSRPFLNENGAVCSPFKQ 360
| | | | |
DB 301 VASHMMVNMNGYTTGVGEQELRGGHPSGDFGRASMDWLRRSSRPFLNENGAVCSPFKQ 360
| | | | |
QY 361 HVPGGGSTYVGLERQSLKKSRENDMAQNLQDPAEDTSSRPKSSLKLPKGI LKKS 420
| | | | |
DB 361 HVPGGGSTYVGLERQSLKKSRENDMAQNLQDPAEDTSSRPKSSLKLPKGI LKKS 420
| | | | |
QY 421 TSSGEVQEDPOELRPVDPDTGQVPVAVSLPRKGI LKKSQRESGYSSPESPESGELD 480
| | | | |
DB 421 TSSGEVQEDPOELRPVDPDTGQVPVAVSLPRKGI LKKSQRESGYSSPESPESGELD 480
| | | | |
QY 481 ASDVFVSGDPVEQKSPQASGLLHRKGI LKNGKFSRTALBGTTPSTFGSLDQLASSHPA 540
| | | | |
DB 481 ASDVFVSGDPVEQKSPQASGLLHRKGI LKNGKFSRTALBGTTPSTFGSLDQLASSHPA 540
| | | | |
QY 541 ARSPRSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNLRGLEQPPSEGLKRMWQ 600
| | | | |
DB 541 ARSPRSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNLRGLEQPPSEGLKRMWQ 600
| | | | |
QY 601 ESI LGSDFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 601 ESI LGSDFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |

RESULT 2
US-10-322-281-23
; Sequence 23, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PaateSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-322-281-23

Query Match 96.8%; Score 3186; DB 16; Length 640;
Best Local Similarity 98.3%; Pred. No. 1.1e-199;
Matches 620; Conservative 0; Mismatches 1; Indels 10; Gaps 5;
```

```
DB 192 KLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPFLY -DEIVNGKPFVGVSEVDSWSL 249
| | | | |
QY 241 GULLYILVHGTMFPDGDHKTLYKQISNGAYREPPKSDACGLIRWLMLNPTRRATLED 300
| | | | |
DB 250 GULLYILVHGTMFPDGDHKTLYKQISNGAYREPPKSDACGLIRWLMLNPTRRATLED 309
| | | | |
QY 301 VASHMMVNMNGYTTGVGEQELRGGHPSGDFGRASMDWLRRSSRPFLNENGAVCSPFKQ 360
| | | | |
DB 310 VASHMMVNMNGYTTGVGEQELRGGHPSGDFGRASMDWLRRSSRPFLNENGAVCSPFKQ 369
| | | | |
QY 361 HVPGGGSTYVGLERQSLKKSRENDMAQNLQDPAEDTSSRPKSSLKLPKGI LKKS 420
| | | | |
DB 370 HVPGGGSTYVGLERQSLKKSRENDMAQNLQDPAEDTSSRPKSSLKLPKGI LKKS 429
| | | | |
QY 421 TSSGEVQEDPOELRPVDPDTGQVPVAVSLPRKGI LKKSQRESGYSSPESPESGELD 480
| | | | |
DB 430 TSSGEVQEDPOELRPVDPDTGQVPVAVSLPRKGI LKKSQRESGYSSPESPESGELD 489
| | | | |
QY 481 ASDVFVSGDPVEQKSPQASGLLHRKGI LKNGKFSRTALBGTTPSTFGSLDQLASSHPA 540
| | | | |
DB 490 ASDVFVSGDPVEQKSPQASGLLHRKGI LKNGKFSRTALBGTTPSTFGSLDQLASSHPA 549
| | | | |
QY 541 ARSPRSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNLRGLEQPPSEGLKRMWQ 600
| | | | |
DB 550 ARSPRSGAVSEDSI LSSSEFPDQDLPERLPETPLRGCVSVNLRGLEQPPSEGLKRMWQ 609
| | | | |
QY 601 ESI LGSDFSLTDCQEVTAAYRQALGICSKLS 631
| | | | |
DB 610 ESI LGSDFSLTDCQEVTAAYRQALGICSKLS 640
| | | | |

RESULT 3
US-10-343-514-41
; Sequence 41, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBEVRRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 630
; TYPE: PRT
; ORGANISM: RAT
US-10-343-514-41

Query Match 93.2%; Score 3068.5; DB 16; Length 630;
Best Local Similarity 94.0%; Pred. No. 4.9e-192;
Matches 593; Conservative 12; Mismatches 25; Indels 1; Gaps 1;
```

Db 121 EVFENS SKIYIVMEYASRGDLYDIYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 180
QY 181 KLENILLDANGNIKIDPGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGEVDSWSL 240
Db 181 KLENILLDANGNIKIDPGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGEVDSWSL 240
QY 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLMVNPTRRATLED 300
Db 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLMVNPTRRATLED 300
QY 301 VASHMVMNNGYTTGVGEQALRECGHPGSDPGASADWLRSSRRLLENAGAVCSFFKQ 360
Db 301 VASHMVMNNGYTTGVGEQALRECGHPGSDPGASADWLRSSRRLLENAGAVCSFFKQ 360
QY 361 HYPGGGSTVPGLEERQSLKSKSRKENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
Db 361 HYPGGGSTVPGLEERQSLKSKSRKENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
QY 421 TSSGEVQEDPOELRPVDPTRGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
Db 421 TSSGEVQEDPOELRPVDPTRGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
QY 481 ASDVFSVGPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
Db 481 ASDVFSVGPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
QY 541 ARSRPSGAVSEDSILSSSFQDLPERLPETPLRGCVSVNLRGLQEPSPSG---LK 596
Db 541 ARSRPSGAVSEDSILSSSFQDLPERLPETPLRGCVSVNLRGLQEPSPSG---LK 596
QY 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 601 ESIGDSCFSLTDCQEVTAAYRQALGICSKLS 631

RESULT 4
US-09-963-159-2
; Sequence 2, Application US/09963159
; Patent No. US20020077312A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-50U1
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-159-2

Query Match 85.8%; Score 2824.5; DB 9; Length 628;
Best Local Similarity 85.7%; Pred. No. 4.2e-176;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLOPSPQASALASASARPLADGLIKSPRLMKQAVKHHHKNLRHYEFL 60
Db 1 MESLVARHSGPTPS---AAELARPLABGLIKSPRLMKQAVKHHHKNLRHYEFL 56
QY 61 ETLGKGYGVKVKARSSGRLVAIKIRKQIKDEODLIRREIEIEMSLNPHITAIH 120
Db 57 ETLGKGYGVKVKARSSGRLVAIKIRKQIKDEODLIRREIEIEMSLNPHITAIH 116
QY 121 EVFENS SKIYIVMEYASRGDLYDIYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 180
Db 117 EVFENS SKIYIVMEYASRGDLYDIYISERPLINERDARHFRQIVSAHYCHONGIVHRDL 176

QY 181 KLENILLDANGNIKIDPGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGEVDSWSL 240
Db 177 KLENILLDANGNIKIDPGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGEVDSWSL 236
QY 241 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLMVNPTRRATLED 300
Db 237 GVLVLYLVHGTMPFDQDHTLVKQISNGAYREPPEPSACGLIRWLMVNPTRRATLED 296
QY 301 VASHMVMNNGYTTGVGEQALRECGHPGSDPGASADWLRSSRRLLENAGAVCSFFKQ 360
Db 297 VASHMVMNNGYTTGVGEQALRECGHPGSDPGASADWLRSSRRLLENAGAVCSFFKQ 356
QY 361 HYPGGGSTVPGLEERQSLKSKSRKENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 420
Db 357 HYPGGGSTVPGLEERQSLKSKSRKENDMAONLQDPAEDTSSRPKSLKPKILKKKSS 416
QY 421 TSSGEVQEDPOELRPVDPTRGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
Db 417 TSSGEVQEDPOELRPVDPTRGQVPVAVSLPRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 473
QY 481 ASDVFSVGPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 540
Db 474 ASDVFSVGPVEQKSPQASGLILHRKGIILKNGKFSRTALEGTPSTFGSLDQLASHPA 533
QY 541 ARSRPSGAVSEDSILSSSFQDLPERLPETPLRGCVSVNLRGLQEPSPSG---LK 596
Db 534 ARSRPSGAVSEDSILSSSFQDLPERLPETPLRGCVSVNLRGLQEPSPSG---LK 593
QY 597 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLS 628

RESULT 5
US-10-423-543-44
; Sequence 44, Application US/10423543
; Publication No. US20040058355A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Libermann, Rosana K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Miyoung
; APPLICANT: Williamson, Mark J.
; APPLICANT: Siles-Santiago, Imaculada
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
; FILE REFERENCE: MP103-0230NM
; CURRENT APPLICATION NUMBER: US/10/423,543
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 10/278,036
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 09/711,216
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/205,447
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 10/012,055
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/248,325
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 10/003,690
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/248,893
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/797,039


```
APPLICANT: GANDHI, Aameena R.
APPLICANT: TRIBULEY, Catherine M.
APPLICANT: CHAMLA, Narinder K.
APPLICANT: YAO, Monique G.
APPLICANT: LU, Dyring Aina M.
APPLICANT: GREENWALD, Sara R.
APPLICANT: RAMKOMAR, Jayalaxmi
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: KEARNEY, Liam
APPLICANT: BURFORD, Neil
APPLICANT: NGUYEN, Damien B.
APPLICANT: TRAN, Y. Tom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: HE, Ann
APPLICANT: THORNTON, Michael
APPLICANT: HAPALIA, April
APPLICANT: ARVIZU, Chandra S.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: LO, Terence P.
APPLICANT: KHAH, Farrah A.
APPLICANT: RECIPON, Shirley A.
APPLICANT: AZIMZAI, Yalda
APPLICANT: POLICKY, Jennifer L.
APPLICANT: DING, Li
APPLICANT: GREITHER, Megan
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: BATER, Sajeev
APPLICANT: ISON, Craig H.
APPLICANT: TITLE OF INVENTION: HUMAN KINASES
APPLICANT: FILE REFERENCE: PI-0125 PCT
APPLICANT: CURRENT APPLICATION NUMBER: US/10/311,034
APPLICANT: CURRENT FILING DATE: 2002-12-10
APPLICANT: PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
APPLICANT: 60/228,056
APPLICANT: PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
APPLICANT: 25
APPLICANT: NUMBER OF SEQ ID NOS: 52
APPLICANT: SOFTWARE: PERL Program
APPLICANT: SEQ ID NO 12
APPLICANT: LENGTH: 594
APPLICANT: TYPE: PRT
APPLICANT: ORGANISM: Homo sapiens
APPLICANT: FEATURE:
APPLICANT: NAME/KEY: misc feature
APPLICANT: OTHER INFORMATION: Inocyte ID No. US20040023242A1 4841542CD1
US-10-311-034-12

Query Match      82.6%; Score 2720.5; DB 15; Length 594;
Best Local Similarity 87.3%; Pred. No. 2.4e-169;
Matches 521; Conservative 22; Mismatches 47; Indels 7; Gaps 2;
```

```
DB 301 WLRSSRPILLENGAKVCSFFKQHPGGSSTTGLERQSHLKKSRKENDMAQSHSTPAD 360
QY 399 TSSRPKSLKCPKGLKKKSSSTSGEVOEDPOELRPVDTGPQVPAYSLPRKGLTK 458
DB 361 TAREKSNULPKGLKKKVSASAGVEDDEPESLPIASQOAP---LPPKGLTK 417
QY 459 SRORESGYSSPESSESGELDASDVFSGDPVEOKSPQASGLLRKGLKINGKPSRT 518
DB 418 PRORESGYSSPESSESGELDAGDVFSGDPVEOKSPQASGLLRKGLKINGKPSRT 477
QY 519 ALEGTTSTFGSLDQASHPPARPERPGAVSEDSILSSSEFPDLDLPERLPETLRGC 578
DB 478 ALELAPTTFGSLDLELAPPPLARSRPGAVSEDSILSSSEFPDLDLPERLPETLRGC 537
QY 579 VSDVNLRLGLEOPPSG---LKRMOESLGDSCFSLTDCQVYAAAYRQALGICSKLS 631
DB 538 VSDVNLRLGLEOPPSG---LKRMOESLGDSCFSLTDCQVYAAAYRQALGICSKLS 594

RESULT 8
US-10-322-281-26
; Sequence 26, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-281-26

Query Match      82.4%; Score 2712.5; DB 16; Length 616;
Best Local Similarity 84.3%; Pred. No. 8.5e-169;
Matches 535; Conservative 25; Mismatches 52; Indels 23; Gaps 9;
```

Db 409 ASAAGVQEDPPELSPID--GQAAP---LLPKKGLKKPRQRESGYSSPESESGELD 463
QY 481 ASAVFVSQDPVEKSAQASGLLHKRGIILKNGKFSRTALEGTTPTFGSLDOLASSHPA 540
Db 464 AGGVFSGPKEKPPQASGLLHKRGIILKNGKFSRTALEGTTPTFGSLDOLASSHPA 523
QY 541 AAPRPSGAVSESISSSEFPDLDLPERLPETPLRCGVSDNLRGLRQPPSG---LK 596
Db 524 AARSRSQAVSESISSSEFPDLDLPERLPETPLRCGVSDNLRGLRQPPSGCLR 583
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 584 RWRQDPLGDSFSLTDCQEVTA--YQALRVCSKLT 616
RESULT 9
US-10-363-829-404
; Sequence 404, Application US/10363829
; Publication No. US20040142331A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;
; APPLICANT: Dahl, Christopher R.; Momiyama, Monika G.;
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;
; APPLICANT: Gerstlin, Jr., Edward H.; Peralta, Careyna H.;
; APPLICANT: David, Marie H.; Panzer, Scott R.;
; APPLICANT: Flores, Vincent Z.; Dafio, Abel;
; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;
; APPLICANT: Chang, Simon C.; Au, Alan P.;
; APPLICANT: Imman, Rebekah R.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1183 USN
; CURRENT APPLICATION NUMBER: US/10/363,829
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: PCT/US01/27628
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/229,751
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,749
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,750
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,747
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,748
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,583
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,517
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,610
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,597
; PRIOR FILING DATE: 2000-09-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PERL Program
; SEQ ID NO 404
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: LG:982800.1.orfl:2000SEP08
US-10-363-829-404
Query Match 54.1%; Score 1781.5; DB 16; Length 406;
Best Local Similarity 83.6%; Pred. No. 3.3e-108;

Matches 342; Conservative 16; Mismatches 44; Indels 7; Gaps 2;
QY 227 GRRYVQEVDSWGLVLYLVHGTMPFDGDKTLVKQISNGAYREPPKPSDACGLIRW 286
Db 1 GRRYVQEVDSWGLVLYLVHGTMPFDGDKTLVKQISNGAYREPPKPSDACGLIRW 60
QY 287 LLMVNTTRAAITLEDVASHMMVMWGTYTGGQEQALREGGHPSDFRASMADLRRSSRP 346
Db 61 LLMVNTTRAAITLEDVASHMMVMWGTYTGGQEQALREGGHPSDFRASMADLRRSSRP 120
QY 347 LLENGAKVCSFFKQHPGGGSTVPGLERQHSLLKSKKENDMAONLQGDPAEDTSSRPGKS 406
Db 121 LLENGAKVCSFFKQHPGGGSTVPGLERQHSLLKSKKENDMAONLQGDPAEDTSSRPGKS 180
QY 407 SLKLPKGLIKKKSSTSSGVEQDPELRPVPTPGQVPVAVSLPPKGLIKKSRQRESGY 466
Db 181 NLKLPKGLIKKKSSTSSGVEQDPELRPVPTPGQVPVAVSLPPKGLIKKSRQRESGY 227
QY 467 YSSPESSEGGELLDADVDFVSGDPVQKSPQASGLLHKRGIILKNGKFSRTALEGTT 526
Db 238 YSSPESSEGGELLDADVDFVSGDPVQKSPQASGLLHKRGIILKNGKFSRTALEGTT 297
QY 527 TFGSLDOLASSHPAARSPSGAVSEDSILSSEFPDLDLPERLPETPLRCGVSDNLRG 586
Db 298 TFGSLDOLASSHPAARSPSGAVSEDSILSSEFPDLDLPERLPETPLRCGVSDNLRG 357
QY 587 LKQPPSEG---LKRWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631
Db 358 LKQPPSEG---LKRWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 406
RESULT 10
US-10-363-829-504
; Sequence 504, Application US/10363829
; Publication No. US20040142331A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;
; APPLICANT: Dahl, Christopher R.; Momiyama, Monika G.;
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;
; APPLICANT: Gerstlin, Jr., Edward H.; Peralta, Careyna H.;
; APPLICANT: David, Marie H.; Panzer, Scott R.;
; APPLICANT: Flores, Vincent Z.; Dafio, Abel;
; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;
; APPLICANT: Chang, Simon C.; Au, Alan P.;
; APPLICANT: Imman, Rebekah R.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1183 USN
; CURRENT APPLICATION NUMBER: US/10/363,829
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: PCT/US01/27628
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/229,751
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,749
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,750
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,747
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,748
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,583
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,517
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,610
; PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: US 60/230,597
PRIOR FILING DATE: 2000-09-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 506
SOFTWARE: PERL Program
SEQ ID NO 504
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: LI:1144409.1.orfl:2000SEP08
US-10-363-829-504

Query Match 54.1%; Score 1781.5; DB 16; Length 406;
Best Local Similarity 83.6%; Pred. No. 3,3e-108;
Matches 342; Conservative 16; Mismatches 44; Indels 7; Gaps 2;

QY 227 GKPYGPEVDWSLGLVLYLVHGTMPFDGQDHRKTLVKQISNGAYRPPKPSDAGLIRW 286
DB 1 GKPYGPEVDWSLGLVLYLVHGTMPFDGQDHRKTLVKQISNGAYRPPKPSDAGLIRW 60
QY 287 LLMVNTTRATLEDVASHVWVWNGYTTGVGEQALRGHPSGDFGASADWLRRSSRP 346
DB 61 LLMVNTTRATLEDVASHVWVWNGYTTGVGEQALRGHPSGDFGASADWLRRSSRP 120
QY 347 LLENGKVCSPFQOHVPGGSGTVGLERHSLKSRKENDMAONLQGDPAEDTSSRPGKS 406
DB 121 LLENGKVCSPFQOHVPGGSGTVGLERHSLKSRKENDMAONLQGDPAEDTSSRPGKS 180
QY 407 SLKLPKGLIKKKSSTSGEVDPEQLRPVDTFGQVPAVSLPRKGLIKSRORBSGY 466
DB 181 NLKLPKGLIKKKSSTSGEVDPEQLRPVDTFGQVPAVSLPRKGLIKSRORBSGY 237
QY 467 YSSPESSEGLDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTALBGTTPS 526
DB 238 YSSPESSEGLDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTALBGTTPS 297
QY 527 TFGSLQGLASHPARPSPGAVSEDSILSSSPDLDLPERLPENTPLNGCVSYNDLRG 586
DB 298 TFGSLQGLASHPARPSPGAVSEDSILSSSPDLDLPERLPENTPLNGCVSYNDLRG 357
QY 587 LEOPPESEG---LKRWQBSLGDSCFSLTDCOEVTAAVROALGICSKLS 631
DB 358 LEOPPESEGSCLRMRQDPLGDSCSFSLTDCOEVTAAVROALGICSKLT 406

RESULT 11
US-09-780-949-2
Sequence 2, Application US/09780949

Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Weich, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
TITLE OF INVENTION: Kinase
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 661
TYPE: PRT
ORGANISM: H. sapiens
US-09-780-949-2

Query Match 51.6%; Score 1699.5; DB 9; Length 661;
Best Local Similarity 55.5%; Pred. No. 1.4e-102;
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

QY 14 PSASALASSAPRLADGLIKSPRLKKAQVKRHHKHNIRHYELETIGKGTGKVK 73
DB 13 PDGLGAPSPREAVAGATTALEP-RKPGVKKHHKHNIRHYELETIGKGTGKVK 71
QY 74 ARS-SEGLVATISIKDKIKQEDLHTRREIEMSLNHPHIIHFEVNSKVIY 132
DB 72 ATERFSGRVVAISIKDKIKQEDLHTRREIEMSLNHPHIIHFEVNSKVIY 131
QY 133 MEYASRGDLYDYISERPRISERDARHPFQIVSALYCHONGIYVHDLKLENTLIDANGN 192
DB 132 MEYASRGDLYDYISERPRISERDARHPFQIVSALYCHONGIYVHDLKLENTLIDANGN 191
QY 193 IKIADFGSLNLYKGFELQFCGSPLYASBEIYNGKRYGPEVDWSLGLVLYLVHGTM 252
DB 192 IKIADFGSLNLYKGFELQFCGSPLYASBEIYNGKRYGPEVDWSLGLVLYLVHGTM 251
QY 253 PFDGQDHRKTLVKQISNGAYRPPKPSDAGLIRWLLMVNTTRATLEDVASHVWVWNGY 312
DB 252 PFDGQDHRKTLVKQISNGAYRPPKPSDAGLIRWLLMVNTTRATLEDVASHVWVWNGY 311
QY 313 TGVGEQALRGHPSGDFGASADWLRRSSRPLLENGKVCSPFQOHVPGGSGTVGL 372
DB 312 TGVGEQALRGHPSGDFGASADWLRRSSRPLLENGKVCSPFQOHVPGGSGTVGL 362
QY 373 ERQHSLSKSRKENDMAONLQGDPAEDTSSRPGKSILKPKGLIKKS-----STSGEVQ 427
DB 363 ERQHSLSKSRKENDMAONLQGDPAEDTSSRPGKSILKPKGLIKKS-----STSGEVQ 418
QY 428 -----EDPDEL-----RPVDTFGQVPAVSLPRKGLIKSRORBSGY 466
DB 419 GVVGPALPSTFKMEQDLCRTGVLLPSSPEAEVPGKLSPKQASATMPKGLIKKQOORBSGY 478
QY 467 YSSPESSEGLDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTAL 520
DB 479 YSSPESSEGLDADVDVPSGDPVBOKSPQASGLLHRKGLIKLNGKESRTAL 536
QY 521 EGTPPS---TFGSLQ-----LASHPARPSPGAVSEDSILSSSPDLDLPER 569
DB 537 DPAVSPPEMPTLESLEPGVAPAGLSRSTY-----SRPSSVTSDDSVLSSSFDLDDIQEN 591
QY 570 LP-ETPLRGCVSYNDLRG-----LRLGQPESEGLKRWQBSLGDSCFS-LTDCOEVTAAV 621
DB 592 RPARQRIKSCVSAENFLQIDFGLQNRPPQYIKRYKRNLDSSFSLLTDMDDVTGVYK 651
QY 622 QALGICSKLS 631
DB 652 QALGICSKLN 661

RESULT 12

US-09-780-949-6
Sequence 6, Application US/09780949
Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Weich, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
TITLE OF INVENTION: Kinase
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 661
TYPE: PRT
ORGANISM: H. sapiens
US-09-780-949-6

Query Match 51.6%; Score 1699.5; DB 9; Length 661;
Best Local Similarity 55.5%; Pred. No. 1.4e-102;
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

14 PSASALASESAPRLADGLIKSPKPLMKKQAVKSHHKNLRRHYELETGKGTGKVK 73
13 PDGLGAPGSPREAVAGATALEP-RKPHGVKHHHKNLKRHYELETGKGTGKVK 71
74 ARS-SGRVATIKSKDKIKDEODLHIREIEMSSLNHPHIIAHEVPENSKIVY 132
72 ATERFGRVVAIKSKDKIKDEODVHIREIEMSSLNHPHIIISIVEFENKDKVII 131
133 MEYASGDLYDIYSERPRLSERDARHFRQIVSAHYCHONGIVHRLKLENTILDANGN 192
132 MEYASGELYDIYSERPRLSERDARHFRQIVSAHYCHONGIVHRLKLENTILDANGN 191
193 IKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGKPYVPEVDSMSLGLVLYIHGTM 252
192 IKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGKPYVPEVDSMSLGLVLYIHGTM 251
253 PRDGDHKTLYKQISNGATREPPKPSDACGLIRMLMNVNPTRRATLEDVASHMVMWGYT 312
252 PRDGFPHKXLIHQISSEGEYREPTQPSDARGLIRMLMNVNPTRRATLEDVASHMVMWGYK 311
313 TGVGEOLREGHPSGDFGRASMDWLRSSRPILLENGAKVCSFKQHPVGGSTVPGU 372
312 SSVCDCDLHDSSEPL-LARIIDWHRSGLQADTEAKMKGLAK--PTTSEVM-L 362
373 ERQSLKSKRKENDMAONLQGDPAEDTSRPGKSLKPKGILKKS-----STSGEVQ 427
363 ERQSLKSKRKENDMAONLQGDPAEDTSRPGKSLKPKGILKKS-----STSGEVQ 418
428 -----EDPDEL-----RVPDTFGQVPAVS-LPRKGIKKSRORESEY 466
419 GVVGPALPSTFKMEODLCRTGVLLPSSPEAEVPGKLSPOQSAITMPKGIKKTQORESEY 478
467 YSSPESSESGELLDSDVAVSGDPVPGKSP-----QASGLLHRRGILKLNKSPRTAL 520
479 YSSPESSESGELLDSDVAVSGDPVPGKSP-----QASGLLHRRGILKLNKSPRTAL 536
521 EGTTPS-----TFGSLDQ-----LASHPARPSPSGAVSEDSIISSEFPQDLPER 569
537 DPALVSPEMPTELESLEPPVPAEGLRSY-----SRPSVYISDVSLSFPLLDQEN 591
570 LP-ETPLRGCVSYDN-----LRGLEQPSSEGLKRWQBSLGDSCFS-LTDCQEVTAAYR 621
592 RPARQRIKSCVSNENFLQIDFEGLQNRPRPQYIKRYRNLADSSPSLLTDMDVQYQVK 651
QY 622 QALGICSKXS 631
DB 652 QALGICSKXS 661

RESULT 13
US-10-354-358-82
Sequence 82, Application US/10354358
Publication No. US20030157082A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macbeth, Kyle J.
APPLICANT: Teat, Feng-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33220, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,

TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
FILE REFERENCE: MP102-020PRKONMIM
CURRENT FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 82
LENGTH: 661
TYPE: PRF
ORGANISM: Homo sapiens
US-10-354-358-82

Query Match 51.6%; Score 1699.5; DB 14; Length 661;
Best Local Similarity 55.5%; Pred. No. 1.4e-102;
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

14 PSASALASESAPRLADGLIKSPKPLMKKQAVKSHHKNLRRHYELETGKGTGKVK 73
13 PDGLGAPGSPREAVAGATALEP-RKPHGVKHHHKNLKRHYELETGKGTGKVK 71
74 ARS-SGRVATIKSKDKIKDEODLHIREIEMSSLNHPHIIAHEVPENSKIVY 132
72 ATERFGRVVAIKSKDKIKDEODVHIREIEMSSLNHPHIIISIVEFENKDKVII 131
133 MEYASGDLYDIYSERPRLSERDARHFRQIVSAHYCHONGIVHRLKLENTILDANGN 192
132 MEYASGELYDIYSERPRLSERDARHFRQIVSAHYCHONGIVHRLKLENTILDANGN 191
193 IKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGKPYVPEVDSMSLGLVLYIHGTM 252
192 IKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGKPYVPEVDSMSLGLVLYIHGTM 251
253 PRDGDHKTLYKQISNGATREPPKPSDACGLIRMLMNVNPTRRATLEDVASHMVMWGYT 312
252 PRDGFPHKXLIHQISSEGEYREPTQPSDARGLIRMLMNVNPTRRATLEDVASHMVMWGYK 311
313 TGVGEOLREGHPSGDFGRASMDWLRSSRPILLENGAKVCSFKQHPVGGSTVPGU 372
312 SSVCDCDLHDSSEPL-LARIIDWHRSGLQADTEAKMKGLAK--PTTSEVM-L 362
373 ERQSLKSKRKENDMAONLQGDPAEDTSRPGKSLKPKGILKKS-----STSGEVQ 427
363 ERQSLKSKRKENDMAONLQGDPAEDTSRPGKSLKPKGILKKS-----STSGEVQ 418
428 -----EDPDEL-----RVPDTFGQVPAVS-LPRKGIKKSRORESEY 466
419 GVVGPALPSTFKMEODLCRTGVLLPSSPEAEVPGKLSPOQSAITMPKGIKKTQORESEY 478
467 YSSPESSESGELLDSDVAVSGDPVPGKSP-----QASGLLHRRGILKLNKSPRTAL 520
479 YSSPESSESGELLDSDVAVSGDPVPGKSP-----QASGLLHRRGILKLNKSPRTAL 536

```

QY 521 EGTTPS----TQSLDQ-----LASSHPAARPSRPSGAVSDDTLTSSSPQDLPPER 569
DB 537 DPALVSPENPTLESISEPVPAGLSRSY-----SRPSSVITSDSVLSDSPDLLOEN 591
QY 570 LP-ETPLRGCVSVN-----LRGLEQPSSEGLKRWMOBSLGDSCPS-LTDCOEYTAAR 621
DB 592 RPARQRIRSCVSAENFLQIDPFGQLQNRPRPOLYKRYRLADSSLSLTDMDVTVQYK 651
QY 622 QALGICSKLS 631
DB 652 QALGICSKLN 661

RESULT 14
US-10-343-514-103
; Sequence 103, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEBVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 103
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-514-103

Query Match 40.2%; Score 1325; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 YEFLETLGKGTGKVKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 116
DB 1 YEFLETLGKGTGKVKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 60
QY 117 IAIHEVFNSSKIVIMEYASRGDLVDYISERPLSERDARHFFROIYSALHYCHONGIY 176
DB 61 IAIHEVFNSSKIVIMEYASRGDLVDYISERPLSERDARHFFROIYSALHYCHONGIY 120
QY 177 HBDLKLIENTLLDANGNIKIADFGLSNLVHKGFLOTFGSSPLVASPEIYNGKPYGPEVD 236
DB 121 HBDLKLIENTLLDANGNIKIADFGLSNLVHKGFLOTFGSSPLVASPEIYNGKPYGPEVD 180
QY 237 SMSLGLTILVVGMPFGQDQDKTIYKQISNGAYRPPKPSDACGLIRMLMWNPTRRR 296
DB 181 SMSLGLTILVVGMPFGQDQDKTIYKQISNGAYRPPKPSDACGLIRMLMWNPTRRR 240
QY 297 TLEDVASHMWV 307
DB 241 TLEDVASHMWV 251

RESULT 15
US-09-836-392-20
; Sequence 20, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, a

```

```

; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PTO2021
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-392-20

Query Match 39.5%; Score 1301.5; DB 9; Length 530;
Best Local Similarity 52.7%; Pred. No. 1e-76;
Matches 290; Conservative 64; Mismatches 125; Indels 71; Gaps 16;

QY 133 MEYASRGDLVDYISERPLSERDARHFFROIYSALHYCHONGIYVHDLKLENTLLDANGN 192
DB 1 MEYASRGDLVDYISERPLSERDARHFFROIYSALHYCHONGIYVHDLKLENTLLDANGN 60
QY 193 IKIADFGLSNLVHKGFLOTFGSSPLVASPEIYNGKPYGPEVDSLSGLTILVVGHTM 252
DB 61 IKIADFGLSNLVHKGFLOTFGSSPLVASPEIYNGKPYGPEVDSLSGLTILVVGHTM 120
QY 253 PRGQDQDKTIYKQISNGAYRPPKPSDACGLIRMLMWNPTRRATLEDVASHMWVMMGYT 312
DB 121 PRGQDQDKTIYKQISNGAYRPPKPSDACGLIRMLMWNPTRRATLEDVASHMWVMMGYT 180
QY 313 TGVGEGQALREGHPSGDFGRASMDWLRRSSRPILLNENAKYCSFFKQHVPGGQSTVPGI 372
DB 181 TGVGEGQALREGHPSGDFGRASMDWLRRSSRPILLNENAKYCSFFKQHVPGGQSTVPGI 231
QY 373 EROHSLKRSKRNDAONIQSDPAEDTSRPGKSLKPKGILKKKS-----STSGEVQ 427
DB 232 EROHSLKRSKRNDAONIQSDPAEDTSRPGKSLKPKGILKKKS-----STSGEVQ 287
QY 428 -----EDPQEL-----RPVDPFGQPVPAVS-LLPRKGIKKRSORESGY 466
DB 288 GIVGPALEPTFKMEQDLCTGVILPSPFAEYVQKLSPPKQSAATMPKKGILKTKQORESGY 347
QY 467 YSPSPESSEGLLDASDVFSQDQVYQKSP-----QASGLLHRRGILKNGKFSRTAL 520
DB 348 YSPSPESSEGLLDASDVFSQDQVYQKSP-----QASGLLHRRGILKNGKFSRTAL 405
QY 521 EGTTPS----TQSLDQ-----LASSHPAARPSRPSGAVSDDTLTSSSPQDLPPER 569
DB 406 DPALVSPENPTLESISEPVPAGLSRSY-----SRPSSVITSDSVLSDSPDLLOEN 460
QY 570 LP-ETPLRGCVSVN-----LRGLEQPSSEGLKRWMOBSLGDSCPS-LTDCOEYTAAR 621
DB 461 RPARQRIRSCVSAENFLQIDPFGQLQNRPRPOLYKRYRLADSSLSLTDMDVTVQYK 520
QY 622 QALGICSKLS 631
DB 521 QALGICSKLN 530

```

Search completed: May 11, 2005, 14:42:20
 Job time : 132.912 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 22:46:12 ; Search time 1080.14 Seconds
(without alignments)
10726.042 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atgagagtcggtgcctact.....gaatctgtcaagctcagc 1893

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Liering first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/PCr_NEM_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/PCrUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/1/pubpna/US11_NEM_PUB.seq:*
- 21: /cgn2_6/prodata/1/pubpna/US60_NEM_PUB.seq:*
- 22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2902	16	US-10-355-975-4
2	1893	100.0	3073	18	US-10-322-281-22
3	1888.2	99.7	2027	18	US-10-343-514-87
4	1683.4	88.9	2026	18	US-10-343-514-27
5	1683.4	88.9	2929	18	US-10-343-514-1
6	1360.2	71.9	1884	9	US-09-963-159-3
7	1360.2	71.9	1884	17	US-10-423-543-45
8	1360.2	71.9	3353	9	US-09-963-159-1
9	1360.2	71.9	3353	17	US-10-423-543-43
10	1357.6	71.7	3360	17	US-10-311-034-38
11	1357	71.7	3404	18	US-10-322-281-25

12	1357	71.7	3463	18	US-10-618-941-11	Sequence 11, Appl
13	1355.4	71.6	2501	17	US-10-302-172-215	Sequence 215, App
14	1355.4	71.6	3443	18	US-10-370-715B-639	Sequence 639, App
15	1071.4	56.6	1186	18	US-10-343-514-13	Sequence 13, Appl
16	1059.4	56.0	37278	18	US-10-322-281-21	Sequence 21, Appl
17	943.4	49.8	1186	18	US-10-343-514-70	Sequence 70, Appl
18	840.2	44.4	2616	18	US-10-363-829-151	Sequence 151, App
19	840.2	44.4	2619	18	US-10-363-829-250	Sequence 250, App
20	723	38.2	39699	18	US-10-322-281-24	Sequence 24, Appl
21	648	34.2	707	18	US-10-343-514-91	Sequence 91, Appl
22	648	34.2	712	18	US-10-343-514-31	Sequence 31, Appl
23	644.2	34.0	661	18	US-10-343-514-93	Sequence 93, Appl
24	613.6	32.4	741	18	US-10-343-514-32	Sequence 32, Appl
25	592	31.3	734	18	US-10-343-514-92	Sequence 92, Appl
26	583.2	30.8	588	18	US-10-343-514-97	Sequence 97, Appl
27	557.4	29.4	660	18	US-10-343-514-33	Sequence 33, Appl
28	533	28.2	578	18	US-10-343-514-95	Sequence 95, Appl
29	515	27.2	548	18	US-10-343-514-96	Sequence 96, Appl
30	510.6	27.0	579	18	US-10-343-514-35	Sequence 35, Appl
31	506.4	26.8	1986	9	US-09-780-949-3	Sequence 3, Appl1
32	506.4	26.8	2884	9	US-09-780-949-5	Sequence 5, Appl1
33	506.4	26.8	6828	9	US-09-780-949-1	Sequence 1, Appl1
34	506.4	26.8	6828	16	US-10-354-358-81	Sequence 81, Appl
35	506.4	26.8	6828	17	US-10-172-118-1541	Sequence 1541, Ap
36	506.4	26.8	6828	18	US-10-342-887-1541	Sequence 1541, Ap
37	506.4	26.8	6828	18	US-10-723-860-2552	Sequence 2552, Ap
38	506.4	26.8	6854	18	US-10-723-860-6703	Sequence 6703, Ap
39	493.6	26.1	585	18	US-10-343-514-37	Sequence 37, Appl
40	483.8	25.6	548	18	US-10-343-514-36	Sequence 36, Appl
41	448.2	23.7	521	18	US-10-343-514-94	Sequence 94, Appl
42	417.8	22.1	521	18	US-10-343-514-34	Sequence 34, Appl
43	383.4	20.3	439	18	US-10-343-514-10	Sequence 10, Appl
44	381.4	20.1	436	18	US-10-343-514-14	Sequence 14, Appl
45	365.6	19.3	421	18	US-10-343-514-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-10-355-975-4
; Sequence 4, Application US/10355975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2823-A
; CURRENT APPLICATION NUMBER: US/10/355,975
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579,664B
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-355-975-4

Query Match 100.0%; Score 1893; DB 16; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 123 ATGAGAGCGGCGGCGGCTGAGCGGCGCTCAAGTCGCTTAACCTTGATGAAG 120
61 TCGGAGAGCGGCGGCGGCTGAGCGGCGCTCAAGTCGCTTAACCTTGATGAAG 120
|||||

Db 183 TCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTGCCTAAACCTGTATGAG 242
QY 121 AAGCAGCGCGTGAAGCGGCACATCACAAAACAACCTCGCGCACCCGTACAGATTCTTG 180
Db 243 AACACAGCGGTGAAGCGGCACATCACAAAACAACCTCGCGCACCCGTACAGATTCTTG 302
QY 181 GAGACGTGGGCAAGGCACTTACGCGGAGGTGAAGAGGACGAGAGCTGGGGCGT 240
Db 303 GAGACGTGGGCAAGGCACTTACGCGGAGGTGAAGAGGACGAGAGCTGGGGCGT 362
QY 241 CTGGTGGCCATCAAGTCATCAGGAAAACAATAAGATGACAGATCTGCTGCAC 300
Db 363 CTGGTGGCCATCAAGTCATCAGGAAAACAATAAGATGACAGATCTGCTGCAC 422
QY 301 ATACGAGGAGATGATGATCATCTTCACTCAACACCCCACTCATATTCATCAT 360
Db 423 ATACGAGGAGATGATGATCATCTTCACTCAACACCCCACTCATATTCATCAT 482
QY 361 GAAGTGTGAGATACAGCAAGATTGTGATGTGATGATGATGATGATGATGATGAT 420
Db 483 GAAGTGTGAGATACAGCAAGATTGTGATGTGATGATGATGATGATGATGATGAT 542
QY 421 CTGTATGATTAATCAATCAGTGAAGCGGCACAGGCTGATGAGCGGACGATTTCTTC 480
Db 543 CTGTATGATTAATCAATCAGTGAAGCGGCACAGGCTGATGAGCGGACGATTTCTTC 602
QY 481 CGACAGATCGTGTCTGCTGCTGCACTACTGCAACCAAGAGATGATGATGATGAT 540
Db 603 CGACAGATCGTGTCTGCTGCTGCACTACTGCAACCAAGAGATGATGATGATGAT 662
QY 541 AAGCTGGAACCACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 663 AAGCTGGAACCACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 722
QY 601 TCCAACTGTACCAACAAGGCAAGTTCCTCAGACGTTCTGTGGAAGCCCTCTCTACGC 660
Db 723 TCCAACTGTACCAACAAGGCAAGTTCCTCAGACGTTCTGTGGAAGCCCTCTCTACGC 782
QY 661 TCGCTGAGATGATCAACGCGGAAGCCCTATGTGCGCCCAAGATGATGATGATGAT 720
Db 783 TCGCTGAGATGATCAACGCGGAAGCCCTATGTGCGCCCAAGATGATGATGATGAT 842
QY 721 GGGGTTCTCTGTATCATCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 843 GGGGTTCTCTGTATCATCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 902
QY 781 ACACTGTGAGAAATCACTAAACGCGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCC 840
Db 903 ACACTGTGAGAAATCACTAAACGCGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCC 962
QY 841 TGTGCTCTGATCCGCTGCTGTGTATGATGATGATGATGATGATGATGATGATGAT 900
Db 963 TGTGCTCTGATCCGCTGCTGTGTATGATGATGATGATGATGATGATGATGATGAT 1022
QY 901 GTAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1023 GTAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 961 CTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 1083 CTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
QY 1021 CGTCTGCTCTGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1143 CGTCTGCTCTGCGCCCTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 1202
QY 1081 CAGTGTCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1203 CAGTGTCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1262
QY 1141 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1263 TCCGGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322

QY 1201 TCTGCGCTTGGCAAGACAGCTTTAAGCTTCCGAAAGGCAATTCTCAAGAAAAGTCTCT 1260
Db 1323 TCTGCGCTTGGCAAGACAGCTTTAAGCTTCCGAAAGGCAATTCTCAAGAAAAGTCTCT 1382
QY 1261 ACCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1383 ACCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1442
QY 1321 GGGCAGCTGTCTCTGTATCTCTGCTCCAGGAAAGGATCTTTAAGAGTCTGA 1380
Db 1443 GGGCAGCTGTCTCTGTATCTCTGCTCCAGGAAAGGATCTTTAAGAGTCTGA 1502
QY 1381 CAGCTGATCTGTATCTATCTCTCTCCAGAGCCAGGAGATCTGGGAACTTTAGAC 1440
Db 1503 CAGCTGATCTGTATCTATCTCTCTCCAGAGCCAGGAGATCTGGGAACTTTAGAC 1562
QY 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1563 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
QY 1501 CTCTCTCTCCAGCAGGAGGATTTCAAACTCAATGAGCAAGTCTCCGCAAGCTTTA 1560
Db 1623 CTCTCTCTCCAGCAGGAGGATTTCAAACTCAATGAGCAAGTCTCCGCAAGCTTTA 1682
QY 1561 GAAAGCACTACCTCTGACACTTTGCTCTCTGACCAACTGAGCTCTCTCAATCTGCA 1620
Db 1683 GAAAGCACTACCTCTGACACTTTGCTCTCTGACCAACTGAGCTCTCTCAATCTGCA 1742
QY 1621 GCGCGGCGCAGCGCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db 1743 GCGCGGCGCAGCGCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1802
QY 1681 TTTGACCAATTGATCTGTGCTGAAGCTTTCCGAAACCCCACTGAGGAGCTGTGTCT 1740
Db 1803 TTTGACCAATTGATCTGTGCTGAAGCTTTCCGAAACCCCACTGAGGAGCTGTGTCT 1862
QY 1741 GTGACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1863 GTGACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1922
QY 1801 GAATCTGTGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1923 GAATCTGTGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
QY 1861 AGACAAGCCCTAGAAATCTGCTCAAGCTCAGC 1893
Db 1983 AGACAAGCCCTAGAAATCTGCTCAAGCTCAGC 2015

RESULT 2
US-10-322-281-22
; Sequence 22, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; FILE REFERENCE: Novel Compositions and Methods in Cancer
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 3073
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-322-281-22

Query Match 100.0%; Score 1893; DB 18; Length 3073;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGTGGTGGCTTTACTCAGAGCCGAGCCAGGCTCCCTCGGCTCGCCCTGGCC	60
Db	92	ATGAGTGGTGGCTTTACTCAGAGCCGAGCCAGGCTCCCTCGGCTCGCCCTGGCC	151
QY	61	TCGAGAGCCGCCGCGCTGGCGGACGGGCTCATCAAGTCGCTTAACTCTGATGAAG	120
Db	152	TCGAGAGAGCCGCCGCGCTGGCGGAGGGGCTCATCAAGTCGCTTAACTCTGATGAAG	211
QY	121	AAGCAGGGGGTGAAGCGGCACCTACAAACACAACTTGGCGCACCGCTACGAGTTCCTG	180
Db	212	AAGCAGGGGGTGAAGCGGCACCTACAAACACAACTTGGCGCACCGCTACGAGTTCCTG	271
QY	181	GAGACGCTGGGCAAGGGGACCTTACGGGAAGGTGAAGAGGACGAGAGCTCGGGGGCT	240
Db	272	GAGACGCTGGGCAAGGGGACCTTACGGGAAGGTGAAGAGGACGAGAGCTCGGGGGCT	331
QY	241	CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAAAGATGACAGAGATCTGCTGCAC	300
Db	332	CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAAAGATGACAGAGATCTGCTGCAC	391
QY	301	ATACGGAGGGAGATTGAGATCATGTCTTCACTCAACACCCCCACATCATTTGCCATCCAT	360
Db	392	ATACGGAGGGAGATTGAGATCATGTCTTCACTCAACACCCCCACATCATTTGCCATCCAT	451
QY	361	GAAAGTTTGAATAGACAGCAAGATTGATTTGATCAGAGTATGCCAGCGCGAGGCGAT	420
Db	452	GAAAGTTTGAATAGACAGCAAGATTGATTTGATCAGAGTATGCCAGCGCGAGGCGAT	511
QY	421	CTGTATGATTTACATCAGTGAAGCGGCACGGCTGATGAGCGGGAAGCGGACATTTCTTC	480
Db	512	CTGTATGATTTACATCAGTGAAGCGGCACGGCTGATGAGCGGGAAGCGGACATTTCTTC	571
QY	481	CGACAGATCGTGTCTGCCCTGSCACTACCTGCGACAGAAAGGAGTGTTCACCGAGATCTC	540
Db	572	CGACAGATCGTGTCTGCCCTGSCACTACCTGCGACAGAAAGGAGTGTTCACCGAGATCTC	631
QY	541	AAGCTGGAAAACTCCTCTTATGATGCGCAATGGAAACATCAAGATTGTGACTTTGGGCTCG	600
Db	632	AAGCTGGAAAACTCCTCTTATGATGCGCAATGGAAACATCAAGATTGTGACTTTGGGCTCG	691
QY	601	TCCAACCTGTACCAACAAGCAAGTTCCTCAGACGTTCTGTGGAGGCGCTCTCTACGCGC	660
Db	692	TCCAACCTGTACCAACAAGCAAGTTCCTCAGACGTTCTGTGGAGGCGCTCTCTACGCGC	751
QY	661	TCGCGTGAATAGTCAACGGGAAGCCCTATGTGGGCCCAAGAGTGAACAGTGTCTCTG	720
Db	752	TCGCGTGAATAGTCAACGGGAAGCCCTATGTGGGCCCAAGAGTGAACAGTGTGTCTG	811
QY	721	GGCGTCTCTCTGTACATCTCTGTGTGATGGCACCATGCGCTTTGACGGGCAAGATCATTAA	780
Db	812	GGCGTCTCTCTGTACATCTCTGTGTGATGGCACCATGCGCTTTGACGGGCAAGATCATTAA	871
QY	781	AACCTGGGAAGCAATAGTAAAGGGGCTTAAAGGTGAGCGGCCCAACCGGTCCGATGCC	840
Db	872	AACCTGGGAAGCAATAGTAAAGGGGCTTAAAGGTGAGCGGCCCAACCGGTCCGATGCC	931
QY	841	TGTGGCCCTGATCCGGTGGCTGTTATGTGTGAACCCACCCGTGGGGCAACTGGAGGAT	900
Db	932	TGTGGCCCTGATCCGGTGGCTGTTATGTGTGAACCCACCCGTGGGGCAACTGGAGGAT	991
QY	901	GTAGCCAGTCAATTGGTGGGTCACTGGGGTTTACACACCGGAGTGTGGGGAAACGGAAGCC	960
Db	992	GTAGCCAGTCAATTGGTGGGTCACTGGGGTTTACACACCGGAGTGTGGGGAAACGGAAGCC	1051
QY	961	CTGGGTGAAGGGTGGGCAACCTAGTGTGTACTTTGGCGGGGCTCCATGGCGGACTGTGTTA	1020
Db	1052	CTGGGTGAAGGGTGGGCAACCTAGTGTGTACTTTGGCGGGGCTCCATGGCGGACTGTGTTA	1111
QY	1021	CGTGGCTCTCGCGGCCCTCTCTGAGATGAGCCAAAGTGTGACGCTTTCAAGCAG	1080
Db	1112	CGTGGCTCTCGCGGCCCTCTCTCTGAGATGAGCCAAAGTGTGACGCTTTCAAGCAG	1171
QY	1081	CACGTGCCGGAGGTGAAGCACTGTACTTGGGCTGAGCGGCAACTTCTCTTAAGAG	1140

Db	1172	CAGGTGCGGGAGGTGAGAGCATGTACTGTGGCTGAGGGGCAACATTCTTTAAGAG	1231
Qy	1141	TCCCGAAAGGAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGATCTCT	1200
Db	1232	TCCGAAAGGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGATCTCT	1291
Qy	1201	TCCTGGCCCTGGCAAGACAGCGCTTAACCTCCGAAAGGACATTCGAAAAAGTCTCT	1260
Db	1292	TCCTGGCCCTGGCAAGACAGCGCTTAACCTCCGAAAGGACATTCGAAAAAGTCTCT	1351
Qy	1261	ACCTGCTCAGGGGAGGTACAGAGAGACCTCAGGAACTCAGACCGGTGCTGATCTCA	1320
Db	1352	ACCTGCTCAGGGGAGGTACAGAGAGACCTCAGGAACTCAGACCGGTGCTGATCTCA	1411
Qy	1321	GGGCAAGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAGCATCTTAAGAGTCTCA	1380
Db	1412	GGGCAAGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAGCATCTTAAGAGTCTCA	1471
Qy	1381	CAGGTGAATCTGGTTTACTACTCTCTCCAGAGCCCGACGAGCTGGGAAATCTTGAC	1440
Db	1472	CAGGTGAATCTGGTTTACTACTCTCTCCAGAGCCCGACGAGCTGGGAAATCTTGAC	1531
Qy	1441	GCCAGTATGTTGTTGAGGTGGGAGCCCGGTGAGAGAAAGTCTCAACAGCTTCAGG	1500
Db	1532	GCCAGTATGTTGTTGAGGTGGGAGCCCGGTGAGAGAAAGTCTCAACAGCTTCAGG	1591
Qy	1501	CTCTCTCTCCACCGCAAGGGGATCTTCAAACTCAATGGCAAGTCTCCCGCAACGCTTA	1560
Db	1592	CTCTCTCTCCACCGCAAGGGGATCTTCAAACTCAATGGCAAGTCTCCCGCAACGCTTA	1651
Qy	1561	GAGGCACTACCCCTAGACACTTTGGGCTCCCTGGACCAACTGGCTCTCCCATCTGCA	1620
Db	1652	GAGGCACTACCCCTAGACACTTTGGGCTCCCTGGACCAACTGGCTCTCCCATCTGCA	1711
Qy	1621	GCCCGGCGCAAGCGCGCCCTCAGAGGGGCTGTAGAGGACAGCATCCTGCTCCGATCC	1680
Db	1712	GCCCGGCGCAAGCGCGCCCTCAGAGGGGCTGTAGAGGACAGCATCCTGCTCCGATCC	1771
Qy	1681	TTTGACCAATTTGACTTGTCTGTAACTCTTCCGAAACCCCATGAGGGGCTGTGTCT	1740
Db	1772	TTTGACCAATTTGACTTGTCTGTAACTCTTCCGAAACCCCATGAGGGGCTGTGTCT	1831
Qy	1741	GTGCAACAACCTGAGGGGCTTTGAGCAGCTTCCTCAGAGGTCGTGAAGCATGTGGCAG	1800
Db	1832	GTGCAACAACCTGAGGGGCTTTGAGCAGCTTCCTCAGAGGTCGTGAAGCATGTGGCAG	1891
Qy	1801	GAATTCCTTGGGGGATAGCTGCTTTTCTGTACAGACTGCCAAGAGGATCGACGCTAC	1860
Db	1892	GAATTCCTTGGGGGATAGCTGCTTTTCTGTACAGACTGCCAAGAGGATCGACGCTAC	1951
Qy	1861	AGACAAGCCCTAGGAATCTGCTCAAACTCAGC	1893
Db	1952	AGACAAGCCCTAGGAATCTGCTCAAACTCAGC	1984
RESULT 3			
US-10-343-514-87			
Sequence 87, Application US/10343514			
Publication No. US20040132025A1			
GENERAL INFORMATION:			
APPLICANT: DROCKER, Daniel J.			
APPLICANT: ROSEN, Cheryl F.			
APPLICANT: LEEBEYRE, Diana L.			
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK			
FILE REFERENCE: DPA-DRUG/PCIT			
CURRENT APPLICATION NUMBER: US/10/343, 514			
CURRENT FILING DATE: 2003-01-31			
PRIOR APPLICATION NUMBER: PCT/CA01/01109			
PRIOR FILING DATE: 2001-08-02			
PRIOR APPLICATION NUMBER: US 60/222, 650			
PRIOR FILING DATE: 2000-08-03			
PRIOR APPLICATION NUMBER: US 60/274, 613			

;; PRIOR FILING DATE: 2001-03-12
;; PRIOR APPLICATION NUMBER: CA 2,340,780
;; PRIOR FILING DATE: 2001-03-28
;; NUMBER OF SEQ ID NOS: 109
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 87
;; LENGTH: 2027
;; TYPE: DNA
;; ORGANISM: MOUSE
US-10-343-514-87

Query Match 99.7%; Score 1888.2; DB 18; Length 2027;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1890; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGATCGGTGGCTTAACTTCAAGCGCCCGAGCGAGGCTCCCTGGGCTCCGCTGAGCC 60
DB 75 ATGAGATCGGTGGCTTAACTTCAAGCGCCCGAGCGAGGCTCCCTGGGCTCCGCTGAGCC 134
QY 61 TCGGAGAGCGCCCGGCTGGCGGAGCGGCTCATCAAGTGCCTAAACCTTGATGAG 120
DB 135 TCGGAGAGCGCCCGGCTGGCGGAGCGGCTCATCAAGTGCCTAAACCTTGATGAG 194
QY 121 AAGCAGCGGTGAAGCGGACCATCAAAACAACCTGCGGACCGCTAGAGTTCTG 180
DB 195 AAGCAGCGGTGAAGCGGACCATCAAAACAACCTGCGGACCGCTAGAGTTCTG 254
QY 181 GAGACGCTGGGCAAGGCACTTAAGGAGGTGAAGAGGAGGAGAGGCTCGGGGCT 240
DB 255 GAGACGCTGGGCAAGGCACTTAAGGAGGTGAAGAGGAGGAGAGGCTCGGGGCT 314
QY 241 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGAGGATCTGCTGAC 300
DB 315 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGAGGATCTGCTGAC 374
QY 301 ATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATATTCATCCATCAT 360
DB 375 ATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATATTCATCCATCAT 434
QY 361 GAAGTCTTTGAGATTAAGCAGCAAGATTTGTATTTGATGAGTATTCAGGCGGAGGAT 420
DB 435 GAAGTCTTTGAGATTAAGCAGCAAGATTTGTATTTGATGAGTATTCAGGCGGAGGAT 494
QY 421 CTGATGATTTACATCAGTGAAGCGGCAAGGCTGAGTGAAGGAGGAGGATTTCTTC 480
DB 495 CTGATGATTTACATCAGTGAAGCGGCAAGGCTGAGTGAAGGAGGAGGATTTCTTC 554
QY 481 CGACAGATCGTGTCTGCTGCACTACTGCGCACAGAACGGGATGTTTCAACGAGATCTC 540
DB 555 CGACAGATCGTGTCTGCTGCACTACTGCGCACAGAACGGGATGTTTCAACGAGATCTC 614
QY 541 AAGCTGAAAAATCTCTTCTAGATGCAATGAAAAATCAAGATTGCTGACTTTGGCCTC 600
DB 615 AAGCTGAAAAATCTCTTCTAGATGCAATGAAAAATCAAGATTGCTGACTTTGGCCTC 674
QY 601 TCCAACTGTACCAAAAGGAAAGTTCTCCAGAGGTTCTGTGGAGGCGCTCTTAAGCC 660
DB 675 TCCAACTGTACCAAAAGGAAAGTTCTCCAGAGGTTCTGTGGAGGCGCTCTTAAGCC 734
QY 661 TCGCTGAGATGATCAACGAGAAAGCCTATGTGGGCCCAAGAGTGAACAGTGTCTCTG 720
DB 735 TCGCTGAGATGATCAACGAGAAAGCCTATGTGGGCCCAAGAGTGAACAGTGTCTCTG 794
QY 721 GGGGTTTCTCTGATCATCTGTGTGATGAGCAACATGCCCTTTTGAAGGAGATCATPAA 780
DB 795 GGGGTTTCTCTGATCATCTGTGTGATGAGCAACATGCCCTTTTGAAGGAGATCATPAA 854
QY 781 ACACTGTGAAGCAATCATGTAACGGGGCTTAACGTTGAGCGGCGCCAGCGCTCGAGGCC 840
DB 855 ACACTGTGAAGCAATCATGTAACGGGGCTTAACGTTGAGCGGCGCCAGCGCTCGAGGCC 914
QY 841 TGTGGCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGGTCGGGCAACTGAGAGAT 900

DB 915 TGTGGCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGGTCGGGCAACTGAGAGAT 974
QY 901 GTAGCCAGTATTTGGGGGCACTGGGGTTTACACCAACCGGATCGGGGAAACAGGAAGCC 960
DB 975 GTAGCCAGTATTTGGGGGCACTGGGGTTTACACCAACCGGATCGGGGAAACAGGAAGCC 1034
QY 961 CTGCTGAGAGGTGGGCACTCTAGTGTGATCTTTGGCGGGGCTTCATGCGGAGTGTGA 1020
DB 1035 CTGCTGAGAGGTGGGCACTCTAGTGTGATCTTTGGCGGGGCTTCATGCGGAGTGTGA 1094
QY 1021 GGTGCTCTCGGCGCCCTCTGTGAGATGAGAGCCAGGTTGTGACGCTTCTTCAAGCAG 1080
DB 1095 GGTGCTCTCGGCGCCCTCTGTGAGATGAGAGCCAGGTTGTGACGCTTCTTCAAGCAG 1154
QY 1081 CAGTGTCCGGAAGGTGAGACATCTTACTGTGGCTGGAAGGCGCAACTTCTTTAAGAG 1140
DB 1155 CAGTGTCCGGAAGGTGAGACATCTTACTGTGGCTGGAAGGCGCAACTTCTTTAAGAG 1214
QY 1141 TCCGAAAGGAGATGATCATGCTCAAAATCTGCAAGGTGACCCGCTGAGAGTACTCT 1200
DB 1215 TCCGAAAGGAGATGATCATGCTCAAAATCTGCAAGGTGACCCGCTGAGAGTACTCT 1274
QY 1201 TCTGCGCTTGGCAAGAGCAGCCTTAAGCTTTCGAAAGGCAATTCAGAAAAAGTCTCT 1260
DB 1275 TCTGCGCTTGGCAAGAGCAGCCTTAAGCTTTCGAAAGGCAATTCAGAAAAAGTCTCT 1334
QY 1261 ACTCTCTCAAGGAGATGACAGAGAGACCTTCAAGAACTCAGACCGGTGCTGATCTCCA 1320
DB 1335 ACTCTCTCAAGGAGATGACAGAGAGACCTTCAAGAACTCAGACCGGTGCTGATCTCCA 1394
QY 1321 GGGCAGGCTCTCCCTGTGATCCCTGCTCCCAAGGAAAGGCACTTTAAGAGTCTCGA 1380
DB 1395 GGGCAGGCTCTCCCTGTGATCCCTGCTCCCAAGGAAAGGCACTTTAAGAGTCTCGA 1454
QY 1381 CAGGTGAATCTGATTAATTAATCTCTCTCAAGAGCCAGAGCTTGGGAACTTTAGAC 1440
DB 1455 CAGGTGAATCTGATTAATTAATCTCTCTCAAGAGCCAGAGCTTGGGAACTTTAGAC 1514
QY 1441 GCCAGTATGTTGTTGATGAGGAGACCCCGTGAAGAGTCTCCACAGGCTTCAGGG 1500
DB 1515 GCCAGTATGTTGTTGATGAGGAGACCCCGTGAAGAGTCTCCACAGGCTTCAGGG 1574
QY 1501 CTCTCTCTCCAGGCAAGGCACTTCTCAAACTGAATGGAAGTTCTCCGCAAGCCTTA 1560
DB 1575 CTCTCTCTCCAGGCAAGGCACTTCTCAAACTGAATGGAAGTTCTCCGCAAGCCTTA 1634
QY 1561 GAAAGCACTAACCCCTAGACATTTTGGCTCCCTGAGACCAACTGCTCTCCATCTGCA 1620
DB 1635 GAAAGCACTAACCCCTAGACATTTTGGCTCCCTGAGACCAACTGCTCTCCATCTGCA 1694
QY 1621 GCCCGGCGGAGCGGCTCTGAGGGGCTGTGATGAGAGACAGATCTGTCTCTCGAGTCC 1680
DB 1695 GCCCGGCGGAGCGGCTCTGAGGGGCTGTGATGAGAGACAGATCTGTCTCTCGAGTCC 1754
QY 1681 TTTGACCAATTGACCTTGAACGTTCCCGGAAACCCCACTGAGGGGCTGTGTCT 1740
DB 1755 TTTGACCAATTGACCTTGAACGTTCCCGGAAACCCCACTGAGGGGCTGTGTCT 1814
QY 1741 GTGAGCAACTGAGGGGCTTTGAGACGCTCCCTCAGAAAGTCTGAAGCGATGTGACAG 1800
DB 1815 GTGAGCAACTGAGGGGCTTTGAGACGCTCCCTCAGAAAGTCTGAAGCGATGTGACAG 1874
QY 1801 GAATCTTTGGGGATGAGCTCTTTTCTGTGACAGACTGCCAAGAGTGAAGTGAAGCTTAC 1860
DB 1875 GAATCTTTGGGGATGAGCTCTTTTCTGTGACAGACTGCCAAGAGTGAAGTGAAGCTTAC 1934
QY 1861 AGACAAGCCTTAGAATCTGCTCAAGCTAGC 1893
DB 1935 AGACAAGCCTTAGAATCTGCTCAAGCTAGC 1967

RESULT 4
US-10-343-514-27

Sequence 27, Application US/10343514
Publication No. US20040132025A1
GENERAL INFORMATION:
APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEEBEYRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DBA-DRUCK/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
PRIOR FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 2026
TYPE: DNA
ORGANISM: RAT
US-10-343-514-27

Query Match 88.9%; Score 1683.4; DB 18; Length 2026;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 1 ATGAGTGGTGGCTTACTTCAAGCCGCCGAGCGAGGCTCCCTGGGCTCCGCCCTGACC 60
DB 75 ATGAGTGGTGGCTTACTTCAAGCCGCCGAGCGAGGCTCCCTGGGCTCCGCCCTGACC 134
QY 61 TCGGAGGCGCCGCCGCCCTGCGGAGCGGCTCATCAAGTGCCTTAACCTCTATGAAG 120
DB 135 ACGGAGGCGCCGCCGCCCTGCGGAGCGGCTCATCAAGTGCCTTAACCTCTATGAAG 194
QY 121 AAGCAGGCGGTGAAGCGGCACATCACAAAACAACCTGCGGACCGCTAGAGTTCTG 180
DB 195 AAGCAGGCGGTGAAGCGGCACATCACAAAACAACCTGAGGACCGCTAGAGTTCTG 254
QY 181 GAGACGCTGGGCAAGGSCACCTTACGGGAGGTGAAGAGGCAAGAGAGCTCGGGGCT 240
DB 255 GAGACGCTGGGCAAGGSCACCTTACGGGAGGTGAAGAGGCAAGAGAGCTCGGGGCT 314
QY 241 CTGTGGCCATCAAGTCTCATGAGAAAGCAAAATCAAGATGACAGATCTCTGAC 300
DB 315 CTGTGGCCATCAAGTCTCATGAGAAAGCAAAATCAAGATGACAGATCTCTGAC 374
QY 301 ATACGGAGGAGATTGAGATGATCTTCACTCAACACCCGCCCATCATTTGCCATCAT 360
DB 375 ATACGGAGGAGATTGAGATGATCTTCACTCAACACCCGCCCATCATTTGCCATCAT 434
QY 361 GAAGTGTGGAATAGCAGCAGAGATTGTGATGTCAAGTATGACAGCGGAGCAT 420
DB 435 GAAGTGTGGAATAGCAGCAGAGATTGTGATGTCAAGTATGACAGCGGAGCAT 494
QY 421 CTGTATATTAATCATCAAGTGAAGCGGCACCGCTGAGTGAAGCGGACCGGAGCATTTCTTC 480
DB 495 CTGTATATTAATCATCAAGTGAAGCGGCACCGCTGAGTGAAGCGGAGCATTTCTTC 554
QY 481 CGACAGATCGTGTGCTGCGCTGACCTAATGCGCAGAGAGCGGAGTGTTCACCGAGATCTC 540
DB 555 CGACAGATCGTGTGCTGCGCTGACCTAATGCGCAGAGAGCGGAGTGTTCACCGAGATCTC 614
QY 541 AAGCTGAAAAACATCTTCTAGATGCAATGAAACATCAAGATGTGATCTTGGCTC 600
DB 615 AAGCTGAAAAACATCTTCTAGATGCAATGAAACATCAAGATGTGATCTTGGCTC 674
QY 601 TCCAACTGTATCAAAAGGCAAGTTCTTCAAGAGTTCTGTGGAGGCTCTCTAGCC 660
DB 675 TCCAACTGTATCAAAAGGCAAGTTCTTCAAGAGTTCTGTGGAGGCTCTCTAGCC 734

QY 661 TCGCTTGAATAGTCAACGGGAGCCCTATGTGGGCCCAAGGTGACAGCTGTCTGTG 720
DB 735 TCACTTGAATAGTCAACGGGAGCCCTATGTGGGCCCAAGGTGACAGCTGTCTGTG 794
QY 721 GGGTGTCTCTGATCAATCTGAGTGAAGGACCAATGCGCTTTGACGGGAGGATCATATA 780
DB 795 GGGTGTCTCTGATCAATCTGAGTGAAGGACCAATGCGCTTTGACGGGAGGATCATATA 854
QY 781 ACACTGTGAAGCAATCAATAGAGGAGGCTTACCGTAGCGGACCGCTCCGATGCC 840
DB 855 ACCCTGTGAAGCAATCAATAGAGGAGGCTTACCGTAGCGGACCGCTCCGATGCC 914
QY 841 TGTGGCTGTATCCGTGAGCTGTATATGATGAACCCCAACCGTGGGCTCACACTGAGAT 900
DB 915 TGTGGCTGTATCCGTGAGCTGTATATGATGAATCCATCGTGGGCTCACACTGAGAT 974
QY 901 GTAGCCAGTATGATGTGGTCAACTGGGGTTACCAACCGGAGTGGGGAACAGAAACC 960
DB 975 GTAGCCAGTATGATGTGGTCAACTGGGGTTACCAACCGGAGTGGGGAACAGAAACC 1034
QY 961 CTGCGTGAAGGATGGGACCCCTAGTGTGATCTTGGCCGGGCTCCATGAGCGGATGTTA 1020
DB 1035 CTGCGTGAAGGATGGGACCCCTAGTGTGATCTTGGCCGGGCTCCATGAGCGGATGTTA 1094
QY 1021 CGTGCCTCTGCGGCCCTCTCTGAGAAATGAGACCAAGGTGSCAGCTTCTCAAGCAG 1080
DB 1095 CGTGCCTCTGCGGCCCTCTCTGAGAAATGAGACCAAGGTGSCAGCTTCTCAAGCAG 1154
QY 1081 CAGCTGCCGGAAGGTGAAGACATCTTCTGTGGGTGAGCGGCAACATTTCTTAAGAG 1140
DB 1155 CAGCTGCCGGAAGGTGAAGACATCTTCTGTGGGTGAGCGGCAACATTTCTTAAGAG 1214
QY 1141 TCCGGAAGGAGATGACATGAGCTCAAAATCTGCAAGGTGACCGGGCTGAGATACCTCT 1200
DB 1215 TCCGGAAGGAGATGACATGAGCTCAAAATCTGCAAGGTGACCGGGCTGAGATACCTCT 1274
QY 1201 TCTGCCCTGAGCAAGACAGCTTAACTTCCGAAAGGCAATTTCAAGAAATGCTCTCT 1260
DB 1275 TCTGCCCTGAGCAAGACAGCTTAACTTCCGAAAGGCAATTTCAAGAAATGCTCTCT 1334
QY 1261 ACTCTGTCAGGAGGATGACAGAGACCTCTCAGAACTCAGACCGGTGCTGATATCTCA 1320
DB 1335 CCTCATCTGGGGGAGTACAGAGGGGCTCTCAGAACTCAGACCGGTGCTGATATCTCA 1394
QY 1321 GGGCAGCTGTCTCTGTATCTCTGCTCCAGAGAAAGGCAATCTTAAGATCTGCA 1380
DB 1395 GGGCAGCTGTCTCTGTATCTCTGCTCCAGAGAAAGGCAATCTTAAGATCTGCG 1454
QY 1381 CAGCGTAATCTGTATCTACTCTCTCAGAGGCCAGGAGTCTGGGAACTCTTAGAC 1440
DB 1455 CAGCGTAATCTGTATCTACTCTCTCAGAGGCCAGGAGTCTGGGAACTCTTAGAC 1514
QY 1441 GCCAGTATGTGTTGTGAGTGGGACCCCGTGAAGAGAGTCTTCAAGAGCTTCAAGG 1500
DB 1515 GCAAGTATGTGTTGTGAGTGGGACCCCGTGAAGAGAGTCTTCAAGAGCTTCAAGG 1574
QY 1501 CTCTCTCTTCAACCGGAGGATCTCAAACTCAATGCAAGTCTCCCGACAGCTTCA 1560
DB 1575 ---CGCTTCAATGCAAGGAGGATCTCAAACTCAATGCAAGTCTCCCGACAGCTTCA 1631
QY 1561 GAAGGCACTACCCCTAGACCTTTGCTCCCTGAGCAACTGAGCTCTCCATCTGCA 1620
DB 1632 GAAGGCACTACCCCTAGACCTTTGCTCCCTGAGCAACTGAGCTCTCCATCTGCA 1691
QY 1621 GCCCGGACCAAGCGGCTCTCAAGGAGTGTGAGTGAAGACAGATCTGTCTCCGAGTCC 1680
DB 1692 GCCCGGACCAAGCGGCTCTCCGAGAGTGTGAGTGAAGACAGATCTGTCTCCGAGTCC 1751
QY 1681 TTTGACCAATTTGACCTTGGCTGAACTGTCTCCGAAACCCCACTGAGAGGCTGTGTCT 1740
DB 1752 TTTGACCAATTTGACCTTGGCTGAACTGTCTCCGAAACCCCACTGAGAGGCTGTGTCT 1811

Db 1583 --- GGCCTCCATCGGAGGCGATCTCAAATCTCAATGCGAAGTTTTCGGACAGCTTGA 1639

Qy 1561 GAAGGCACTACCCCTAGACCTTTGGCTCCCTGAGCCAATGCGCTCTCTCCATCTTGA 1620

Db 1640 GAAAGCACTGCGCCCTAGACCACTTTGGCTCCCTAGACCAACTGCGCTCTCTCATCTTACA 1639

Qy 1621 GCCCGGCGCAAGCGCCGCTCTAGAGGGCTGTGAGTGAAGGACAGCATCTGTCTCTCCAGTCC 1680

Db 1700 GCCCGGGGCAAGCGCTCTCTCGGAGACTGTGAGTGAAGGACAGCATCTGTCTCTCCAGATCC 1759

Qy 1661 TTTGACCAATTGGAAGTTGCTGTGAACGTCTTCCCGAACCCTCAGAGGGCTGTGTGTCT 1740

Db 1760 TTTGACCAATTGGAAGTTGCTCCCGAGGCGCTTCCCGAACCCTCAGAGGAGCTGTGTGTCT 1819

Qy 1741 GTGGAACAACCTGAGAGGGGCTGTGAGAGCTCTCCACGAAGTCTGAACCATGTGTGGAG 1800

Db 1820 GTGGAACAACCTGAGAGGCTGTGAGAGCTCTCCACGAAGGCTTAAACGATGTGTGGAG 1879

Qy 1801 GAATCTCTGGGGGATAGCGTCTTTCTCTGACAGATCGGCCAAGAGGTGACTCGAGCCCTAC 1860

Db 1880 GAATCTCTGGGGGATAGCGCTTTCTCTGACAGATCGGCCAAGAGGTGACAGAGCCCTAC 1939

Qy 1861 AGACAAGCCCTAGGAATCTGCTCAAGGCTCAGC 1893

Db 1940 AGACAAGCCCTAGGAATCTGCTCGAAGGCTCAGC 1972

RESULT 6
 US-09-963-159-3
 : Sequence 3, Application US/09963159
 : Patent No. US20020077312A1
 : GENERAL INFORMATION:
 : APPLICANT: CURTIS, RORY A.J.
 : APPLICANT: GALVIN, Katherine M.
 : TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
 : FILE REFERENCE: 101/47-5001
 : CURRENT APPLICATION NUMBER: US/09/963,159
 : CURRENT FILING DATE: 2001-09-25
 : PRIOR APPLICATION NUMBER: US 60/234,922
 : PRIOR FILING DATE: 2000-09-25
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 3
 : LENGTH: 1884
 :
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-963-159-3

Query Match	71.9%;	Score 1360.2;	DB 9;	length 1884;
Best Local Similarity	84.2%;	Pred. No. 0;		
Matches 1565;	Conservative	0;	Mismatches 273;	Indels 21;
				Gaps 2;

QY 47 CCTCCGCGCCTTGACTCTCGGAGAGACGCCCGCGCTGCGGAGACGGAGCTCATCAAGTGGCCTA 106

Db 35 CCACTCCTCGGCGCGCAGAGGTAGCCCGGCGCTGGCGGAGAGGGCTGTATCAAGTGCCTCA 94

QY 107 AACCTCTGATGAAGAGAGCGGGGTGAAGCGGCACATCAAAACAACAACCTGGCGGACAC 166

Db 95 AGCCCTTAATGAGAGAGCAGGGCGGTGAAGCGGCACACACACAGACAACTTGGCGGACAC 154

QY 167 GCTACGAGATTCTTGAGAGCGCTGGGCAAGGACCTACCGGAGGTTGAAGAAAGGACGAG 226

Db 155 GCTACGAGATTCTTGAGAGCCCTGGGCAAAAGGACCTACCGGAGGTTGAAGAAAGGCGCGG 214

QY 227 AGAGCTCGGGGCGTGTGGTGGCCATCAAGTCATCAGAAAGACAAAATCAAAAGATGAGC 286

Db 215 AGAGCTCGGGGCGCTGTGGCCATCAAGTCAATCCGGAGGACAAAATCAAAAGATGAGC 274

QY 287 AGGATCTCTGCACATACGAGGGGAGATTGAGATCATATGTCCTCACTCAACCAACCCGCCACA 346

Db 275 AAGATCTATGCAACATACGAGGGGAGATTGAGATCATATCACTCAACCAACCCCTTACA 334

OY	347	CAATTGCATCATGAAGTGTGTAAGATATGACACAAAGATTGTATGTCATGGAAGTATG	406
Db	335	TCATTGCCATCATGAAGTGTGTAAGAACACACACCAAGATGTATGTCATGGAAGTATG	394
OY	407	CCAGCCGAGGCGATCTGTATGATTAATCAATCAGTAGCGGCGACCGCTGAGTGGCGGACG	466
Db	395	CCAGCCGGGGGACCTTTATGACTACATCAAGCGAGCGGACGACCTGAGTGGCGGAG	454
OY	467	CCAGGCATTTCTTCCGACAGATGCTGTGCTCTGCACTACTGCCACAGAACGGGATCG	526
Db	455	CTAGGCAATTTCTTCCGGCAGATGCTCTCTGCGCGTCACTATGGCCATCAGAACAGATTG	514
OY	527	TTGACCGAGATCTCAAGCTGGAAAAACATCCCTTCAAGTGGCCAAAGGAAACATCAAGATTG	586
Db	515	TCACCCGAGATCTCAAGCTGGAAGAACTCTTGTGATGGCCAAAGGAATATCAAGATTG	574
OY	587	CTGACTTTGGCCTCTCCAACTGTACCAACAAGGCAAGTTCTCTCAGACGTTCTGTGGGA	646
Db	575	CTGACTTTCGGCTCTCCAACTCTACATCAAGGCAAGTTCTCTCAGACATTTGTGGGA	634
OY	647	GCCCTCTCTACGCTTCGCTGAGATGATCAACGGAAAGCCCTATGTGGGCCAGAGGTGG	706
Db	635	GCCCCCTCTATGCTTCGCCAGAGATGTGTCAATGGGAAGCCCTTACACAGGCCCAAGGTTGG	694
OY	707	ACACTGTCCTCTGGGGGCTTCTCCTGTACATCTCTGTGTGACATGGAACATATGCCCTTTGACG	766
Db	695	ACACTGTCCTCTGGGGGCTTCTCCTGTACATCTCTGTGTGACATGGAACATATGCCCTTTGATG	754
OY	767	GGCAGATTCATTAACAAGCTGGTGAAGGAATAATCAATTAACGGGGGCTTACCGTGAAGCCGCCA	826
Db	755	GGCATTAACATTAATAATCTTATGTAAACAGATACAGAACGGGGCTTACCGGAGCCACTTA	814
OY	827	AGCCGTCCGATGCTGTGGCCTGTATCCGATGGCTGTTAATGTGTAAACCCACCCGTGGG	886
Db	815	AACTCTGATAGCTGTGGCCTGTATCCGATGGCTGTTGATGTGTAAACCCACCCGGCGGG	874
OY	887	CCACACTGGAAGATATGACAGTATTTGTGTGGTTCACCTGGGGTTTACCAACCGGAGTGC	946
Db	875	CCACCTTGAAGGAATGTGGCCAGTCACTGTGTGGGTTCACCTGGGGTTTACCGCACCCGAGTGG	934
OY	947	GGAAACAGAAAGCCCTGCTGATGGGTGGGACCTTATGTGTGACTTTTGGCGGGGCTTCA	1006
Db	935	GAGAGCAGAGAGGCTTCGCAATGAGGTGGGACCTTGGCAGTGAATCTTGCCTGGCTTCA	994
OY	1007	TGGCGGACTGTGTAAGTGCCTCTCGCGCCCTCTCTGGAAGATGAGCCAAAGGTGTGCA	1066
Db	995	TGGCTGAATGCTCGCGGCTTCCCTCCGCCCTCTCTGGAAGATGAGGCGCAAGGTGTGCA	1054
OY	1067	GCTTCTTCAAGACGACGATGCGGGGAGGTGGAAGCACTGATCTTGGCTGTGAGCGGCAC	1126
Db	1055	GCTTCTTCAAGACGACGACCTGTGTGGGGAGACCAACCCCTGGCTGTGAGCGGCACG	1114
OY	1127	ATTCTCTTAAGAATCCCGAAAGGAGATGACATGGCTCAAAATCTGCAAGTGAACCCGG	1186
Db	1115	ATTGCTCAAGAAATCCCGAAGGAGATGACATGGCCCAAGTCTTCTCAAGTGAACCGG	1174
OY	1187	CTGAGATTAACCTCTTCTCGCCCTGGCAAGACAGCCTTAAAGCTTCCGAAAGGCAATTTCTCA	1246
Db	1175	CTGATGACACTGCGCCCATGCGCTTGGCAAGACCACTCAAGCTGCCAAAGGCAATTTCTCA	1234
OY	1247	AGAAAAAGTCTCTACTCTGTCAGGGGAGGTACAGAGAGACCTTCAGAACTCAGACCGG	1306
Db	1235	AGAAAGAGGTGTCAAGCTCTGCAAGAAAGGGGTACAGAGAGACCTTCCGAGCTCAGCCCAA	1294
OY	1307	TGCTGTATCTCAAGGGGAGCGTGTCCCTGTGTATCCCTGTCTCCAAAGGAABAGCAATCC	1366
Db	1295	TCCCTGTGAGCCCAAGGAGGCTGTGCC-----CCCTGTCTCCCAAGAGGGCAATTC	1345
OY	1367	TTAAGAAGTCTCAGACAGTGAATCTGTGTTACTACTCTCTCCAGAGCCCAAGCAAGTCTG	1426
Db	1346	TCAAGAGAGCCCGCAAGAGGAGTGTGGCTATCTACTCTCTCCGAGACCCAGTGAATCTG	1405
OY	1427	GGGAATCTTAAAGACCGCAAGTGAATGTGTTTGTGAGTGGGAGACCCCTGTGAAGCAAGTCTC	1486

Db 851 |ACAGCTGTCCTTGGGTTCTCTCTACATCTGCTGATGCAATGCAATGCTTTGATG 910
QY 767 |GGGAGATCAATAAACACTGGTGAAGCAATCACTAAAGGGGCTTAACCTGAGCGCCCA 826
Db 911 |GGGATGACCAATAAGATCTAGTGAACAGATCAGCAAGGGGGCTTACCGGGAGCACCTA 970
QY 827 |AGCGTCGCAATGCTGCTGCTGATCCGGTGGTCTGTTAATGTTGAACCCCACTGGG 886
Db 971 |AACCTCTGATGCTGCTGCTGCTGATCGGTGGCTGTTGATGTTGAACCCCACTGGG 1030
QY 887 |CCAACACTGAGATGTAGCAAGTATGATGATGATCACTGGGGTTAACAACAACCGAGTGC 946
Db 1031 |CAACCTGAGAGATGTAGCAAGTATGATGATGATCACTGGGGTTAACAACAACCGAGTGC 1090
QY 947 |GGGAAACGAAAGCCTGCGTGAAGGGTGGGCACTTAATGATGATGATGATGATGATGAT 1006
Db 1091 |GAGAGCAGAGAGCTCCGCAATGAGGGTGGGCACTTGGGAGTGAATGCTGCGCGGCTTCA 1150
QY 1007 |TGGCGGATGCTGATACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
Db 1151 |TGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1210
QY 1067 |GCTTCTTCAAGACAGCACTGCGCGGAGGTGGAACATGTAACCTGGGCTGAGGCGGAC 1126
Db 1211 |GCTTCTTCAAGACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
QY 1127 |ATTCTCTTAAAGATCCCGAAAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGG 1186
Db 1271 |ATTGCTCAAGAGATCCCGCAAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGG 1330
QY 1187 |CTGAGATACCTCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
Db 1331 |CTGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390
QY 1247 |AGAAAAGATCCTCTACCTGCTCAAGGGAGGTGACAGAGAACCTCAGAACTCAGACGG 1306
Db 1391 |AGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
QY 1307 |TGCCTGATCTCCAGGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366
Db 1451 |TCCCTGAGACCCAGGGGCAAGGCTGCC-----CCTGCTCCCAAGAAAGGCACTTC 1501
QY 1367 |TTTAAAGATCTCCAGACGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426
Db 1502 |TCAAGAGCCCGCAACAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
QY 1427 |GGGAACTCTAGAGCGCAGTGTATGTTGAGTGGGAGCCCGGTGAGAGAGAGCTC 1486
Db 1562 |GGAGGCTCTTGAAGCGAGCGAGCTGTTTGTGATGGGATCCCAAGAGAGAGAGCTC 1621
QY 1487 |CAGAGGCTTGAAGGCTCTCTCCACCGCAAGGCACTTCTCAAACTCAATGCAAGTCT 1546
Db 1622 |CGAAGCTTCAAGGCTCTCTCTCCATGCAAAAGCACTCTCAAACTCAATGCAAGTCT 1681
QY 1547 |CCGCAACAGCTTGAAGAGCACTACCTCTAGCACTTGTGCTGCTGCTGCTGCTGCTGCT 1606
Db 1682 |CCGAGAGAGCTTGAAGAGCTGCGGCGCCCAACCTTGGGCTCTGATGAACTCGCC 1741
QY 1607 |CCTCCATCTGAGCGCGGCGCGAGCGCGCTCAGGGGCTGAGAGAGAGAGAGAGCTC 1666
Db 1742 |CACTGCGCGCGCGCGCGAGCGCGAGCGCGCTTGAAGGCGCTGAGAGAGAGAGAGCTC 1801
QY 1667 |TGTCTCCAGATCTTGAACAAATTGACTTGAAGCTTCTCCGAAACCCCACTGA 1726
Db 1802 |TGTCTCTGAGTCTTGAACAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1861
QY 1727 |GGGCGT 1780
Db 1862 |GGGCGT 1921
QY 1781 |-----GTGAGAGCAATGT 1834

Db 1922 |GAACTGCTGAGGCGCTGCGGAGAGATCTTTGGGGGAGACAGTCTTTTCCCTGACAG 1981
QY 1835 |ACTGCCAAGAGTGAAGTGAAGCTTACAGACAGCCCTTGAAGATCTGCTCAAGCTCAGC 1893
Db 1982 |ACTGCCAAGAGTGAAGTGAAGCTTACAGACAGCCCTTGAAGATCTGCTCAAGCTCAGC 2040
RESULT 9
US-10-423-543
Sequence 43, Application US/10423543
Publication No. US20040058355A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Liberman, Rosana K.
APPLICANT: Hunter, John J.
APPLICANT: Meyers, Rachel E.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Olandt, Peter J.
APPLICANT: Tsai, Fong-ying
APPLICANT: Galvin, Katherine M.
APPLICANT: Chun, Miyoung
APPLICANT: Williams, Mark J.
APPLICANT: Silos-Santiago, Immaculada
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 56638,
TITLE OF INVENTION: 18610, 33117, 21967, h1983, m1983, 38555 OR 593 MOLECULES
FILE REFERENCE: MP103-0230NMIM
CURRENT APPLICATION NUMBER: US/10/423, 543
CURRENT FILING DATE: 2003-04-25
PRIOR APPLICATION NUMBER: US 10/278, 036
PRIOR FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: US 09/711, 216
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/205, 447
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 10/012, 055
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/248, 325
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 10/003, 690
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/248, 893
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/797, 039
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186, 061
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 10/217, 168
Remining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 119
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 3353
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(3353)
OTHER INFORMATION: n = A,T,C or G
US-10-423-543-43
Query Match 71.9%; Score 1360.2; DB 17; Length 3353;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;
QY 47 |CTCCGCGCTGCTGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 106
Db 191 |CAACTCCCTGCGCGAGAGCTAGCCCGCGCTGCGAGAGGAGCTGATCAAGTCCGCCA 250


```

; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAPALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: DING, Li
; APPLICANT: GREYER, Megan
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: BATRA, Sajeev
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
; CURRENT APPLICATION NUMBER: US/10/311,034
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
; 60/228,056
; PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
; 25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PERL Program
; SEQ ID NO 38
; LENGTH: 3360
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CB1
US-10-311-034-38

```

Query Match 71.7%; Score 1357.6; DB 17; Length 3360;
 Best Local Similarity 83.7%; Pred. No. 0;
 Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;

```

QY 32 GCCAGGCTCCCTGGGCTCCGCTGGGCTTGGAGAGCCGCCGCGCTGGCGGACGGCG 91
DB 98 GCGCGGGGCTCCGCGCCCACTCCCTCGCGGAGAGTACCGCGCCGCGGAGAGGGC 157
QY 92 TCATCAAGTGGCTTAACCTCTGATGAAAGACGGGGGTGAAGCGGCACATACAAAC 151
DB 158 TGATCAAGTGGCCCAAGCCCTTAATGAAGACGGGGGTGAAGCGGCACACCAAGC 217
QY 152 ACAACCTGCGGACCGCTAGAGAGTCTGGAGACGCTGGGCAAGGAGGACCTAAGGGAGG 211
DB 218 ACAACCTGCGGACCGCTAGAGAGTCTGGAGACGCTGGGCAAGGAGGACCTAAGGGAGG 277
QY 212 TGAAGAGGACGAGAGAGCTGGGGGCTGTGTGGCCATCAAGTCCATCAGAAAGACA 271
DB 278 TGAAGAGGCGGAGAGAGCTGGGGGCTGTGTGGCCATCAAGTCCATCAGAAAGACA 337
QY 272 AATCAAGATGAGAGAGATCTGTGCAATACGAGAGGAGATTGAATCATGTTCTTAC 331
DB 338 AATCAAGATGAGAGAGATCTGTGCAATACGAGAGGAGATTGAATCATGTTCTTAC 397
QY 332 TCAACCAACCCCATCATTCATGCAATGATGAAGTGTGAAGATGAGAGCAAGTTGTA 391
DB 398 TCAACCAACCTCATCATTCATTCATTCATGAGATGTTGAGAACAGACGAGAGTCTGA 457
QY 392 TTGTCAATGAGATATGCGAGCGAGCGATCTGTATGATTAATCATCATGAGCGGACGGC 451
DB 458 TCCTCAATGAGATATGCGAGCGGAGCGACCTTATATGATCATCATGAGAGGAGAGAGC 517
QY 452 TGATGAGCGGAGCGGAGGAGATTTCTTCCGACGAGATCTGTCTGCGCTGACTACTGCC 511
DB 518 TCATGAGCGGAGGAGGAGATTTCTTCCGACGAGATCTGTCTGCGCTGACTACTATGCC 577
QY 512 ACCAGAACGGGATCTGTTACCGAGATCTCAAGCTGGAAGAAATCTCTTGAATGAGCAAG 571

```

```

DB 578 ATCAAGACAGATTGTCCACCGAGATCTCAAGCTGAGAAACATCTCTTGGGTCCATG 637
QY 572 GAAACATCAAGATTGTGACTTTGGCTCTTCAACCTGTACCAAGAGGAGTTCTCC 631
DB 638 GGAATATCAAGATTGTGACTTGGCTCTTCAACCTGTACCAAGAGGAGTTCTCC 697
QY 632 AGAGTTCTGTGGAGGCTCTTCAAGGCTCTGTAGATAGTCAAGGAGGCTATG 691
DB 698 AGACATCTGTGGAGGCTCTTCAAGGCTCTGTAGATAGTCAAGGAGGCTATCA 757
QY 692 TGGGCCGAGAGTGAAGCTGTGCTTGGAGCTTCTGTATCACTGTGTGATGACA 751
DB 758 CAGGCCAGAGAGTGAAGAGCTGTGCTTGGAGCTTCTGTATCACTGTGTGATGACA 817
QY 752 CCATGCCCTTTGAAGAGGAGAGATCATTAACACTGTGTAAGCAATCAAGTAAAGCGGCT 811
DB 818 CCAATGCCCTTTGAATGAGGATGACATTAAGATCCATGTAAGCAAGTCAAGAGGAGCT 877
QY 812 ACCGTGAGCGGCGCAAGCGCTCCGATGCTGTGTGCTGTATCCGTTGCTTTATGTTGA 871
DB 878 ACCGGAGGACCACTTAACCTCTGTATGCTGTGTGCTGTATCCGTTGCTTTATGTTGA 937
QY 872 ACCCCACCCGTGGGCGCACACTGAGAGATGAGCACTGATGAGGTGAGTCACTGGAGTT 931
DB 938 ACCCCACCCGCGCGGCGCACCTGAGAGATGAGCACTGATGAGGTGAGTCACTGGAGTT 997
QY 932 ACAACACCGAGATGCGGAGAAACAGAAAGCTGTGAGAGGTGAGCACTTATGTTGACT 991
DB 998 ACAGCACCGAGATGAGGAGAGACAGAGGCTCCGATGAGGAGTGGGCACTTGGAGTACT 1057
QY 992 TTGGCCGGGCTCAATGGCGGATCTGTATGTTGCTCTTGGCGCTCTTGGAGATG 1051
DB 1058 CTGCGCGGCTCTCAATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGAGATG 1117
QY 1052 GAGCAAGATGTGCAAGCTTCTTCAAGAGACAGCTGCGGAGGAGTGAAGACGTATCTTG 1111
DB 1118 GCGCCAAAGTGTGCAAGCTTCTTCAAGAGACAGCTGCTGTGAGGAGAGACACCTCTG 1177
QY 1112 GGTGAGCGGCAACATTTCTTAAGAGTCCGAAAGAAATGACATGCTCAAAATC 1171
DB 1178 GCTGAGAGCGCAAGATTTGCTCAAGAAAGTCCGAAAGAAATGACATGAGGAGTCTC 1237
QY 1172 TGCAAGTGAACCGGCTGAGAGATCTTCTTCTGCTGTGCGCAAGAGAGCTTAACTTC 1231
DB 1238 TCCACAGTGAACAGGCTGAGAGACATGCTCCATGCTGTGCAAGAGACCTCAAGCTGC 1297
QY 1232 CGAAAGCATTTCAAGAAAGAGTCTTACCTGTGAGGAGGATCAGAGAGACCTC 1291
DB 1298 CAAGGACATTTCAAGAAAGAGTGTGCTGTGCAAGAGGAGGATCAGAGAGACCTC 1357
QY 1292 AGGAATCAAGACCGTGTGCTGTATCTCAGGAGAGCTGTCTCTGTATCTCTGTCTC 1351
DB 1358 CGGAGCTCAACCCATTCCTGTGAGAGCGGAGGAGGCTGTCTCTCTGTCTCTC 1408
QY 1352 CAAGAAAGCATCTTAAGAGTCTTGAAGTCTGACAGGCTGAATGTGTTACTTCTCTCAG 1411
DB 1409 CCAAGAGGCGCATTTCAAGAAAGCTTCAAGAGGCGGAGTGTGCTACTTCTCTCTCCG 1468
QY 1412 AGCCACGCGATCTGGGGAATCTTTGAAGCGCAAGTATGTTTGAAGTGGGAGCCCG 1471
DB 1469 AGCCACGCGATCTGGGAGCTTTGAGCAGAGCGAGCGTGTGTTGAGTGGGAGTCCA 1528
QY 1472 TGAAGAGAGATCTCAAGGCTTCAAGGCTCTCTCTCAACCGCAAGGAGGATCTCAAC 1531
DB 1529 AGGAGGAGAAAGCTCTCGCAAGCTTCAAGGCTGTCTCTCTCAATCCCAAGAGATCTCAAC 1588
QY 1532 TCAATGCAAGTTCTCCGACAGCTTGAAGGCACTACCTTGAACCTTTGCTGCC 1591
DB 1589 TCAATGCAAGTTCTCCGACAGCTTGAAGGCTGTGAGGCTTCCGAGGCTTCCGAGCTGCC 1648
QY 1592 TGAACCACTGTGCTCTCCATCTGAGGCGGCGGCGGCGGCTCTCAAGGAGGCTGTGA 1651
DB 1649 TGAATGAATCTGCGCCACCTGCGCTGTGCGGCGGCGGCGGAGCCGACCTCAAGGAGCTGTGA 1708

```

QY	1652	GTGAGAGACAGACATCTGTGTCCTCCAGATCCTTTGACCAATTTGGACTTGCCTGAAGGTTC	1711
Db	1709	GGGAGGACAGACATCTGTGTCCTTGTAGTCTTTGACCAAGCTGACCTTGCTGAACGGCTTC	1768
QY	1712	CCGAAACCCCACTGAGGGGCTGTGTGTCTGTGGACAACTGAGGGGGCTTGAAGACCTTC	1771
Db	1769	CAGACCCCCCACTGCGGGGCTGTGTGTCTGTGGACAACTGACCGGGGCTTGAAGAGCCCC	1828
QY	1772	CCTGAGAG-----GTGGAAGCCATGCTGACAGGAATCTTTGGGGGATACT	1819
Db	1829	CCTCAGAGGGCCCTGGAAGCTGCTGTAGAGCGGTGCGGAGGATCTTTTGGGGGACACT	1888
QY	1820	GCTTTTCTCTGACAGACTGCCAAGAGGTGACTGACCTTAAGACAAAGCCCTAGGAATCT	1879
Db	1889	GCTTTTCCCTGACAGACTGCCAGAGGTGACAGCGACTTAACGACAGGCACTGAGGGCTCT	1948
QY	1880	GCTCAAGCTAGC	1893
Db	1949	GCTCAAGCTACAC	1962

```

RESULT 11
US-10-322-281-25
; Sequence 25, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCES: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 666
; SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
;
; LENGTH: 3404
;
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-25

```

Query Match	71.7%;	Score 1357;	DB 18;	Length 3404;
Best Local Similarity	84.1%;	Pred. NO. 0;		
Matches 1563; Conservative	0;	Mismatches 275;	Indels 21;	Gaps 2

OY	47	CCTCGGCGCTGGCGCTCTGGAGAGCGGCCGGCGCTGGCGGACGGGCTCATCAAGTGGCTA	106
Db	162	CCACTCTCCCTCGGCGCGAGAGCTAGCCCGGCGCTGGCGGAGGGGCTGATCAAGTGGCCCA	221
OY	107	AACCTCTGATGAGAGAGCGGCGGTGAGAGCGGCACCATCACAAAACAACCTGCGGCACC	166
Db	222	AGCCCTCATGAGAGAGCGGCGGTGAGAGCGGCACCAACACAGACACAACCTGCGGCACC	281
OY	167	GCTACGAGTTCCTGGAGACGCTGGGCAAGGGCACTTACGGGAAAGGTGAAGAAAGGACGAG	226
Db	282	GCTACGAGTTCCTGGAGACCCCTGGCGAAAGGCACTTACGGGAAAGGTGAAGAAAGGCGCGAG	341
OY	227	AGAGTCGGGGGCGTGTGGTGGCCATCAAGTCCATCAGAAANAACAAATCAAGATGAGC	286
Db	342	AGAGTCGGGGGCGCTGTGGCCATCAAGTCCATCCGAGAGGACAAATCAAGATGAGC	401
OY	287	AGGATCTCTGCACATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACA	346
Db	482	AAGATCTTATGACATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCTCACA	461
OY	347	TCATTTGCATTCATGAAGTGTTTGAGAAATGACAGCAAGATTGTGATTCATGAGATG	406
Db	462	TCATTTGCATTCATGAAGTGTTTGAGAACAGACAGCAAGATCTGATTCGTATGAGATG	521
OY	407	CCAGCGGAGGCGCATCTGTATGATTATCTACGTAGAGCGGCACACGCTGAGTGAAGCGGAGACG	466
Db	522	CCAGCGGAGGCGCACTTTATGATCTACATCAGGAGCGGAGGAGCTCAAGTGAAGCGCGAAG	581

OY	467	CAAGSANTTTCTTCCGACAGATTCGATTCGCGCATCTACCTGACCAAGAAACGAGATCG	526
Db	582	CTAGGCAATTTCTTCGGCAGATTCGTCCTGCGGTGCATATTTGCCATCAGAAACGATTTG	641
OY	527	TTCAACCGAGATCTCAAGCTGGAAAAACATCCTCTTAGATGCAATGCAATGAACATCAAGATTG	586
Db	642	TTCAACCGAGATCTCAAGCTGGAAAAACATCCTCTTAGATGCAATGCAATGCAATGAACATTTG	701
OY	587	CTGACTTTGAGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCAAGCGTTCTGTGGGA	646
Db	702	CTGACTTCGGGCTCTCCAACTGTACCAATCAAGGCAAGTTCTCTGACAGCAATTTGTGGGA	761
OY	647	GGCCTCTCTAGGCTTCGCTGAGATATGTCAACGGAAAGCCCTATGTGGGCCCAAGATGTG	706
Db	762	GGCCCTCTTAGGCTTCGCAAGATTTGTCAATGTGGAAGCCCTTACACAGGCCCAAGATGTG	821
OY	707	ACAGCTGGTCTGTGGGCGTTCTCCTGTACATCCTGTGATGAGGACCAATGCGCTTTGACG	766
Db	822	ACAGCTGGTCCCTGGGATTTCTCTCTACATCTCTGTGATGAGGACCAATGCGCTTTGATG	881
OY	827	AGCGGTCCGATGCTCTGTGCTGATTCGGGTGCGCTTTATGTGTGAACCCCAACCGCTGGG	886
Db	942	AACCTCTGAAGCGCTGTGGCTGTGATCCGGTGTGTTGATGTGTAAACCCCAACCGCGGG	1001
OY	887	CCACACTGAGAGATGTATGACAGTATGTGTGGGTCAACTGGGGGTTAACCAACCGGAGTGG	946
Db	1002	CCACCTTGAGAGATGTGCGCAGTACTGTGTGGGTCAACTGGGGGTTAGCGCCACCGAGTGG	1061
OY	947	GGAACAGGAAAGCCCTGTGAGGGGTGGGCAACCTTAGTGTGACTTTGGCCGGGCTTCA	1006
Db	1062	GAGAGCAGAGAGGCTCCGATATAGGGTGGGCAACCTTGGAATGTACTGTGCGCGCTTCA	1121
OY	1007	TGGCGAATGTTATCGTGCCTCTTCGCGGCCCTCTTGAGATGAGCCAAAGTGTGCA	1066
Db	1122	TGGCTGATCGGCTCCGGGTTCTTCGCCCTCTTGAGAAATGGGGCCAAAGTGTGCA	1181
OY	1067	GCTTCTTCAAGACACAGTGCCTGGAGGTGGAACACTGTACTGTGGGTGTGAGGGGCAAC	1126
Db	1182	GCTTCTTCAAGACACATGCACTGTGTGGGGAAACAACCTTGCGCTTGAAGGGCAACG	1241
OY	1127	ATTCTCTTAAAGATTCGCCAAAGAGATGATCATGTGCTCAAAATCTGCAMAGSTGACCCGG	1186
Db	1242	ATTGCTCTAAGAAATCCCGCAAGAGATGATCATGTGCCAGTCTCTCACAATGTACACGG	1301
OY	1187	CTGAGATTACTCTTCTTCGCGCTTGCAAGACAGCTTTAAGCTTTCGAAAGGCAATTCGA	1246
Db	1302	CTGATGTACACTGCGCCATTCGCTTGCAAGCAACTCAAGCTCCAAAGGGCAATTCGA	1361
OY	1247	AGAAAAAGTCTCTTAATCTCGTCAAGGAGATATCAGAGAACCTTCAGAACTCAGACCGG	1306
Db	1362	AGAAAGAGTGTACGCTCTGCAAAAGGGGTACAGAGAACCTTCGAGCTCAGGCCAA	1421
OY	1307	TGCTGTAAATCTCAAGGAGGCTGATCCCTGCTGATCCCTGCTCCCAAGAAAGGCAATCC	1366
Db	1422	TCCTCTGGAGCCCAAGGCAAGCTGCC-----CCCTGCTCCCAAGAGGGCAATTC	1472
OY	1367	TTAAGAAATCTCAAGACGCTGATCTGTGTTACTACTCTCTCAAGACCCCAAGCGAGTCTG	1426
Db	1473	TCAGAAAGCCCGCAAGCGGAGTGTGGCTACTCTCTCCGAGGCCAGTGAATCTG	1532
OY	1427	GGGAATCTTATGATGCCAGTATGTGTTTGTGATGGGGACCCCGTGGAGCGAAGATCTC	1486
Db	1533	GGGAGCTCTTGAAGCGCAGGCACTGTTTGTGATGTGGGATCCCAAGAGAGGAAGGCTC	1592
OY	1487	CACAGGCTTCAAGGCTCCTCTCTCAACGCAAGGGCAATTCACAACTCAATGTGCAAGTCT	1546
Db	1593	CGCAAGCTTCAAGGCTCCTCTCTCAATGCGCAAGGCAATCTCAAACTCAATGTGCAAGTCT	1652
OY	1547	CCGCGACAGCTTATGAAGGCACTACCCCTAGCACTTTGGCTCCCTGAGCCAACTGGGCT	1606

Db 1653 CCGAGAGAGCTTGGAGACTGGCGGCCCCCACCACCTTGGCTCCCTGGATGAATCTCGCCC 1712
Qy 1607 CTTCCCATCTGACGCGCGCCGACCGCCCTTCAGAGGCTGTGAGTGAAGACAGATCC 1666
Db 1713 CACCTCGCCCTTGGCCCGGCGACGCCACCTCAGAGGCTGTGAGGAGGACGATCC 1772
Qy 1667 TGTCTCCGAGTCTTTTGAACCAATTGACCTTGGCAACGTCCTCCGAAACCCCATCTGA 1726
Db 1773 TGTCTTGAAGTCTTTTGAACCAAGCTTGAACCTTGAACGCTCCAGAGCCCATCTGC 1832
Qy 1727 GGGGCTGTGTCTGTGAGACACACTGAGAGGAGCTTGAAGACCTCCCTCAGAG----- 1780
Db 1833 GGGGCTGTGTCTGTGAGACACACTGAGAGGAGCTTGAAGAGCCCTCAGAGGAGCTGTG 1892
Qy 1781 -----GTGGAAGCATGTGTGAGAGAAATCTTGGAGATAGCTCTTTTCTGTGACG 1834
Db 1893 GAAGCTCTGAGGCGCTGGGCGGAGGATCTTTGGGGGACAGCTGTCTTCCCTGACAG 1952
Qy 1835 ACTGCCAAGAGTACCTGACAGCTTACAGACACAGCCCTTGAAGATCTGTCTCAAGCTCAGC 1893
Db 1953 ACTGCCAAGAGTACAGCACTTACAGACAGGACCTGAGGGTCTGTCTCAAGCTCAGC 2011

RESULT 12

US-10-618-941-11
; Sequence 11, Application US/10618941
; Publication No. US2004019792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618, 941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395, 632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 3463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-11

Query Match 71.7%; Score 1357; DB 18; Length 3463;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

Qy 47 CTTCCGCTGGCTCGGAGCGCCCGCGCTGCGAGCGGCTCATCAAGTCGCTTA 106
Db 223 CCACTCCCTCGCGCGCAGACTAGCCCGCGCTGCGAGAGGCTGATCAAGTCGCCA 282
Qy 107 AACCTGATGAGAGAGCAGCGGTGAAGCGGCATCAAAACAACAACCTGGGCAAC 166
Db 283 AGCCCTTAATGAGAGAGCAGCGGTGAAGCGGCACCAACAACAACAACCTGGGCAAC 342
Qy 167 GCTACAGATTCTGAGAGCGCTGGGCAAGGCACTTACGGGAAGGTGAAGAGGCAAG 226
Db 343 GCTACAGATTCTGAGAGCGCTGGGCAAGGCACTTACGGGAAGGTGAAGAGGCGGG 402
Qy 227 AGAGCTCGGGGCGCTGTGTGAGCATCAAGTCATCAGAGAAACAAATCAAAAGTACG 286
Db 403 AGAGCTCGGGGCGCTGTGTGAGCATCAAGTCATCAGAGAAACAAATCAAAAGTACG 462
Qy 287 AGAGTCTGTCGACATACGAGAGGAGATGAGATCAAGTCTTCACTCAACCCCAACA 346
Db 463 AAGATCTGATGACATACGAGAGGAGATGAGATCAAGTCTTCACTCAACCCCAACA 522
Qy 347 TCATTCGATCCATGAAAGTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 406
Db 523 TCATTCGATCCATGAAAGTGTGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 582

Qy 407 CAGCCGAGGAGATCTGTATGATTACATCATGAGCGGCCACGCGCTGATGAGCGGACG 466
Db 583 CAGCCGCGGCGAGCTTTATGACTACATGACGAGCGCGAGAGCTGATGAGGCGCAAG 642
Qy 467 CAGGCAATTTCTTCCGACATATGTCGTCGCGCGCACTGTCGCAACAGAGGAGATCG 526
Db 643 CTAGGCAATTTCTTCCGACATATGTCGTCGCGCGCACTGTCGCAACAGAGATG 702
Qy 527 TTCACGAGATCTCAAGCTGAGAAACATCTTCTAGATGCAATGGAATCAAGATTG 586
Db 703 TCACCGAGATCTCAAGCTGAGAAACATCTTCTAGATGCAATGGAATCAAGATTG 762
Qy 587 CTGACTTGGCTCTTCAACCTGTACCAAGAGCAAGTTCTTCAGACGTTCTGTGAG 646
Db 763 CTGACTTGGCTCTTCAACCTGTACCAAGAGCAAGTTCTTCAGACATTTCTGTGAG 822
Qy 647 GCCCTCTTACGCGCTGCGCGAGATGTCAGAGGAAAGCCCTTATGTGGGCGAGAGTGG 706
Db 823 GCCCTCTTATGCTGCGCGAGATGTCAGAGGAAAGCCCTTACAGAGCCGAGAGTGG 882
Qy 707 ACAAGTGTCTTGGGCGTTCCTGTACATCTGTGTGATGAGCAACATGCGCTTGTAGC 766
Db 883 ACAAGTGTCTTGGGCGTTCCTGTACATCTGTGTGATGAGCAACATGCGCTTGTATG 942
Qy 767 GGCAGATCTATAAACACTGTGAGCAATCACTAACGCGGCTTACCTGAGCGGCCCA 826
Db 943 GGCATGACATAAAGATCTTGTGAAACATGATACAGCAACGCGGCTTACCGGAGCACTTA 1002
Qy 827 AGCGTCCGATGCTGCGCTGATCCGCGGCGCTTAAATGTGAAACCCACCGCTGGG 886
Db 1003 AACCTTGTATGCTTGGGCGTATCCGCGGCGCTTGTGATGAGTGAACCCACCGCGGG 1062
Qy 887 CCAACTGAGAGATGTAGCAGATCATTTGTGGTCAACTGAGGCTTACACACCGAGTGC 946
Db 1063 CCACTGAGAGATGTAGCAGATCATTTGTGGTCAACTGAGGCTTACACACCGAGTGC 1122
Qy 947 GGAAGACGAGAGCGCTGCGTGAAGGTGGGCAACCTTGTGATCTTTGGCGGCGCTTCA 1006
Db 1123 GAGAGCAGAGAGCGCTGCGATGAGAGGTGGGCAACCTTGTGATCTTTGGCGGCGCTTCA 1182
Qy 1007 TGGCGGATGCTTAAAGTGTGCTCTCGCGCGCCCTCCGAGAGATGAGAGCGAGTGTGCA 1066
Db 1183 TGGCTGATGCTTAAAGTGTGCTCTCGCGCGCCCTCCGAGAGATGAGAGCGAGTGTGCA 1242
Qy 1067 GCTTCTTCAAGACAGATGTCGCGGAGGTGGAAGCACTTACCTGGGCTGAGCGGCAAC 1126
Db 1243 GCTTCTTCAAGACAGATGTCGCGGAGGTGGAAGCACTTACCTGGGCTGAGCGGCAAC 1302
Qy 1127 ATTCTCTTAAGAGTCCCGAAGAGATGATGCTCAAAATCTGCAAGGTGACCGG 1186
Db 1303 ATTCTCTTAAGAGTCCCGAAGAGATGATGCTCAAAATCTGCAAGGTGACCGG 1362
Qy 1187 CTGAGGATACCTTCTTCCGCGCTGCGAGAGCGCTTAAAGCTTCCGAAAGGATTTCTCA 1246
Db 1363 CTGAGGATACCTTCTTCCGCGCTGCGAGAGCGCTTAAAGCTTCCGAAAGGATTTCTCA 1422
Qy 1247 AGAAAAAGTCTTACTCGTCAAGGAGGTGATCAGAGAGAGCTTACAGAACTGACACGG 1306
Db 1423 AGAAAAAGTCTTACTCGTCAAGGAGGTGATCAGAGAGAGCTTACAGAACTGACACGG 1482
Qy 1307 TGCCTGATCTTCCAGGAGCGCTGTCTCGCTGTATCTCTGCTCCAGAGAAAGCATTC 1366
Db 1483 TGCCTGATCTTCCAGGAGCGCGAGGCTGCGC-----CCCTGCTCCCAAGAGGAGCTTC 1533
Qy 1367 TTAAGAGTCTTCCAGGAGCGGTGATCTGTGTTACTTCTTCCAGAGCGCGAGAGTCTG 1426
Db 1534 TCAAGAGCGCGCAAGCGCGGATCTGTGCTTACTTCTTCCAGCGCGAGAGTCTG 1593
Qy 1427 GGAAGCTCTTGAAGCGCAAGTGTGTGTGTGAGTGGGAGCGCGTGGAGCAGAGATCTC 1486
Db 1594 GGAAGCTCTTGAAGCGCGAGGAGTGTGTGTGTGAGTGGGATCTCCAGAGAGCAGAGCTTC 1653

QY 1487 CATAGGCTTGAAGGCTCTCTCTCTGCAAGGCAAGGCAATTTCTCAACTAATGAGCAAGTTCT 1546
DB 1654 CGAAGCTTCAAGGCTGCTCTCTCTCATGCAAAAGGCAATCTCAAACTAATGGCAAGTTCT 1713
QY 1547 CCGGCAAGGCTTGAAGGCACTACCCCTAGCACTTTGGCTCCCTGAGCAACCACTGGGCT 1606
DB 1714 CCGAGACAGCTTGGAGCTGCGGCCCCCACACACCTTGGCTCCCTGAGCAATCTGGCCC 1773
QY 1607 CTTCCCATCTGCAAGCCGCGCCAGCCGCTCTGAGGGGCTGTGATGAGGACAGCATCC 1666
DB 1774 CACTGGCCCCCTGGGCCCGGGCCAGCCGACCTTAGGGGGCTGTGAGGAGGACAGCATCC 1833
QY 1667 TGTCTCCGAGTCTTTGACCAATTGACCTTGGCTAAAGCTCTTCCGAAACCCCACTGA 1726
DB 1834 TGTCTCTGAGTCTTTGACCAAGCTTGGCTAAAGCTCTTCCGAGCCCACTGAC 1893
QY 1727 GGGGCTGTGTCTGTGAGCAACCTGAGGGGGCTTGAAGCAAGCTCTCTCAAGAG----- 1780
DB 1894 GGGGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTGAAGAGCCCTCTCAAGAGGGCCCTG 1953
QY 1781 -----GTCTGAAGCAATGAGTGGAGGATCTTGGGGGATGAGTGTCTTTCTCTGACAG 1834
DB 1954 GAAGCTGCTGAGGCGCTGGCGGCGAGATCTTTGGGGAGACGCTGCTTTCTCTGACAG 2013
QY 1835 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAAGCAAGCCCTAGAAATCTGCTCAAGCTCAGC 1893
DB 2014 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAAGCAAGCAAGTGAAGGCTGTGCTCAAGCTCAGC 2072

RESULT 13

US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dmanac, Radolje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids at
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNDP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc FL_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 71.6%; Score 1355.4; DB 17; Length 2501;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CTTCCGCTTGGCTCTGGAAGCGCCGCGCTGTGGGAGAGGGCTCATCAATGCGCTTA 106
DB 140 CCACTCCCTCGGCGGAGAGCTACCCGCGCTGTGGGAGAGGGCTCATCAATGCGCTTA 199
QY 107 AACCTGATGAAGAACAAGCGGTGAAGGGGCACTACAAACACAACTGGGGACC 166
DB 200 AGCCCTTAATGAAAGAGCGGTGAAGGGGCACTACAAACACAACTGGGGACC 259
QY 167 GCTACGAGTCTCTGAGACGCTGGGCAAGGCACTTACGGGAGGTGAAGAGGACGAG 226

DB 260 GCTACGAGTCTCTGAGACCTTGGGCAAGGCACTTACGGGAGGTGAAGAGGCGCGG 319
QY 227 AGAGCTGGGGGCGCTGTGTGGCAATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 286
DB 320 AGAGCTGGGGGCGCTGTGTGGCAATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 379
QY 287 AGAGTCTGTGACATACGAGGAGATTTGAGATCATGTCTTCACTCAACCACTCCCA 346
DB 380 AAGATCTGATGCAATACGAGGAGATTTGAGATCATGTCTTCACTCAACCACTCCCA 439
QY 347 TCATGTCATCAATGAAGTGTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 406
DB 440 TCATGTCATCAATGAAGTGTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 499
QY 407 CCAAGCCAGAGCGATCTGTATGATTAATCAATGATGAGCGGCAAGCTGATGAGCGGAGC 466
DB 500 CCAAGCCAGAGCGATCTGTATGATTAATCAATGATGAGCGGCAAGCTGATGAGCGGAGC 559
QY 467 CCAAGCATTTCTTCCAGAGATGTGTGCTGCTGCTGCACTACTGCTCACCAGACGAGATCG 526
DB 560 CTAGGCAATTTCTTCCGAGATGTGTGCTGCTGCTGCACTACTGCTCACCAGACGAGATCG 619
QY 527 TTCAACGAGATCTCAAGCTGGAACATCTTCTGATGCAATGAGCAATGAGCAATGAGATG 586
DB 620 TCAACCGAGATCTCAAGCTGGAACATCTTCTGATGCAATGAGCAATGAGCAATGAGATG 679
QY 587 CTGACTTTGGCTCTTCCAACTGTACCAAGGCAAGTCTTCCAGAGCTTCTGTGGGA 646
DB 680 CTGACTTTGGCTCTTCCAACTGTACCAAGGCAAGTCTTCCAGAGCTTCTGTGGGA 739
QY 647 GCCCTCTTACGCTCTGCTGAGATGATCAACGAGAAAGCTTATGTGGGCCAGAGGTGG 706
DB 740 GCCCTCTTACGCTCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 799
QY 707 ACAAGTGTCTCTGTGGGCTTCTCTGTACATCTCTGTGATGATGATGATGATGATGATGATG 766
DB 800 ACAAGTGTCTCTGTGGGCTTCTCTGTACATCTCTGTGATGATGATGATGATGATGATGATG 859
QY 767 GGGAGATCAATTAACACTGATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCA 826
DB 860 GGGAGATCAATTAACACTGATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCA 919
QY 827 AGCGTCCGATGCTGTGGCTGATCCGCTGATCCGCTGATCCGCTGATCCGCTGATCCGCTG 886
DB 920 AACCTCTGATGCTGTGGCTGATCCGCTGATCCGCTGATCCGCTGATCCGCTGATCCGCTG 979
QY 887 CCACACTGAGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 980 CCACCTTGAAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
QY 947 GGGAGCAAGAAAGCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1006
DB 1040 GAGAGCAAGAGGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1099
QY 1007 TGGGGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 1100 TGGGGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159
QY 1067 GCTTCTTCAAGCAAGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126
DB 1220 ATTGCTCAAGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1279
QY 1187 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
DB 1280 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1339
QY 1247 AGAAAAAGTCTTCACTGTGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAG 1306
DB 1340 AGAAAAAGTCTTCACTGTGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAG 1399

Oy	1307	TGCGCTGAATATCCAGAGGAGCGCTGCGCTCGTATTCCTGCTGCCAGAGAAAGGCATCC	1366
Db	1400	TCCCTCGAGAGCCAGAGGACAGGCTGCCCC-----CCTGCTCCCAAGAGGGCATTC	1455
Oy	1367	TTAAGAACTCTGACAGCGTGAATCTGGTTACTACTCTCTTCACAGGCCACAGAGTCTG	1428
Db	1451	TCAAGAAACCCCGACAGCGCGAGTCTGGCTACTACTCTCTCTCCAGCCCACTGAATCTG	1510
Oy	1427	GGAAACTCTTGAAGCGCAGATGATGTTTGATGATGGGAGACCCCTCGAAGAGAAATCTC	1488
Db	1511	GGAGACTTTGGACCCAGGCGACGATGTTTGTATGATGGGATATCCAAAGACAGAAAGCTTC	1570
Oy	1487	CACAGGCTTCAAGGCTCTCTCTCCACCGCAAGGCAATTCTCAAACTCAATGGCAATCTT	1548
Db	1571	CGAAGCTTCAAGGCGTGTCTCTCCATTCGAAAGAGATCTCAAACTCAATGGCAATCTT	1630
Oy	1547	CCCGCACAGCTTGAAGAGCACTACCCCTGAACCTTTGGCTCCTCTGAGCAACTGAGCT	1608
Db	1631	CCAGACACGCTTGAAGCTCGGCGCCCCACACCACTTGGGCTCCTCTGATGAATCGCGCC	1690
Oy	1607	CCTCCCATCTGCAGCGCCGCGCCAGCGCGCCCTCGAGGGGCTGTGATGAGGACAGCATCC	1666
Db	1691	CACCTCGCCCCCTGAGCCCGGCGCAGCCCACTTCAGGGCTGTGTAGCGAGACAGCATCC	1756
Oy	1667	TGTCTCTCGAGTCTTTGACCAATTGACTTGCCTGAACTGTCTTCCGAAACCCCATCTGA	1728
Db	1751	TGTCTCTGAGATCTCTTGTACACAGCTGAGCTTGCCTGAACTGTCTTCCGAACTGTG	1810
Oy	1727	GGGGCTGTGTCTGTGTGACCAACTGAGGGGGCTTGACAGCTCCCTCCAGAAAG-----	1788
Db	1811	GGGGCTGTGTCTGTGTGACCAACTGAGGGGGCTTGAGAGAGCCCCCTCGAAGGGCCCTG	1870
Oy	1781	-----GTCTGAAGGATGTGTGTGAGAGATCTTGGGGGATATGATGCTTTTCTCTGACAG	1834
Db	1871	GAGCTGTGTGTGTGTGTGACCAACTGAGGGGGCTTGAGAGAGCCCCCTCGAAGGGCCCTG	1930
Oy	1835	ACTGCGAAGAGTGACTGACGCTTACAGACAGCCCTTGAAGATCTGCTCAAAAGCTCAGC	1893
Db	1931	ACTGCGAAGAGTGACTGACGCTTACAGACAGGCACTGAGGGGTCTGCTCAAAAGCTCAGC	1989

Query Match	71.6%	Score 1355.4	DB 18	Length 3443
Best Local Similarity	84.0%	Pred. No. 0		
Matches 1562	Conservative 0	Mismatches 276	Indels 21	Gaps 2

47 CCTCGGCTTGAGGCTGAGAGCGCGCGGCTGAGCGAGCTCATCAAGTGGCTTA 106

Db	1127	ATTCTTTAAGAGTCCGGAAGGAAATGACATGCTCAAAATCTGCAGAGTACCCGG	1186
Qy	1182	GCTTCTTCAAGACATGACCTGTGTGGGGAAACAACACCCCTGTGAGGGCCAGC	1241
Db	1067	GCTTCTTCAAGACAGTGCCTGGAGAGTGGAAACATGTACCTGTGGCTGAGCGGCAAC	1126
Qy	1062	GAGAGCAGGAAGCTCCGATAGAGGTGGGACCTCTGCAGTGAACCTTGCCTCCCTCA	1121
Db	1007	TGGGGAGCTGTATACGTGCTCTCCGGGCCCCCTCTGAGAAATGAGACCAAGGTGTGCA	1066
Qy	947	GGGAAACAGGAAGCCCTGTGTGAGGGGTGGGACCTTAAGTGTGACTTTTGGCTGGGCTTCA	1006
Db	887	CCACACTGAGAGATGTAACGACATCTTGTGTGGTCAACTGGGATTACACAACGAGAGTCG	946
Qy	827	AGCGTCGATAGCTGTGGCTGTATCCGGTGTGGTCTGTAAATGTGTAAACCCGACCGTGGG	886
Db	822	ACAGCTGTGCTCTGGGTGTCTCTCTTCAACATCTGTGGTCAATGGCACCATGCTCTTGTATG	881
Qy	767	GGCAGGATTCATAAACAACCTGTGAAAGCAATACATTAACGGGGCTTACCTGTAAAGCCGCCA	826
Db	702	CTGACTTGTGCGCTCTCCAACTGTACCAACAAAGGCAAGTTCTCTCAGACATTTCTGTGGGA	701
Qy	647	GCCCTCTTACGCTCTGCCTGAGATGATCAACGGAAAGCCCTATGTGGGCCACAGAGTGG	706
Db	642	TCACCCGAGATCTCAAGCTGGAACAACATCTCTTGGATGCCAAATGGGAATATCAAGATTG	646
Qy	587	CTGACTTGTGCGCTCTCCAACTGTACCAACAAAGGCAAGTTCTCTCAGACGTTCTGTGGGA	586
Db	527	TTCAACGAGATCTCAAGCTGGAACAAACATCTTCTAGATGCCAATGGAAACATCAAGATTG	526
Qy	522	CCAGGCAATTTCTTCCGACAGATCGTGTCTGCTGCCTGCACCTACCTGCAACCCAGACGGGATCG	521
Db	467	CCAGCCGAGGCGATCTGTATATTAACATCACTGAGCGGCGCACCGCTGTGTAGCGGGAGCG	466
Qy	407	TCATTTGCCATTCATAGAAATGTTTGAATATACAGACAGATTTGTATTTGTCACTGAAGTATG	406
Db	347	TCATTTGCCATTCATAGAAATGTTTGAATATACAGACAGATTTGTATTTGTCACTGAAGTATG	346
Qy	287	AGAGTCTGTGCACATACGAGGAGGAGATTTGATCATGTCTTTCATCTCAACCAACCCCA	286
Db	227	AGAGTCTGTGCAGTCTGTGTGCGCATCAAGTCACTACAGGAAAGACAAATCAAGTATGAC	226
Qy	167	GCTACAGATTCCTGTGAAGCCCTGTGGGCAAGGCACTTACGGGAAGGTGAAGAAAGGCGCGG	166
Db	107	AACCTTGATGAAGACGAGCGGTGAAGCGGCACCATCAAAACAACAACCTTGTGGGCAAC	106
Qy	102	AGCCCTTATGAAGAGCAGCGGTGAAGCGGCACACACAGACACAACTGTGGGCAAC	101
Db	102	AGCCCTTATGAAGAGCAGCGGTGAAGCGGCACACACAGACACAACTGTGGGCAAC	101
Qy	942	AACTCTTGATGCTGTGTGCTGTATCCGGTGTGGTCTGTGATGTGTAAACCCACCGCGGG	941
Db	882	GGCATGACATTAAGATCCTTGTGAAACAAGATCAACCAACGGGGCTTACCGGAGGCACCTTA	881
Qy	822	ACAGCTGTGCTCTGGGTGTCTCTCTTCAACATCTGTGGTCAATGGCACCATGCTCTTGTATG	821
Db	762	GCCCCCTTATAGCTCTCCGCAAGATTTGTCAATGGGAAGCCCTTACAGAGGCCAAGAGTGG	761
Qy	702	CTGACTTGTGCGCTCTCCAACTGTACCAACAAAGGCAAGTTCTCTCAGACATTTCTGTGGGA	701
Db	642	TCACCCGAGATCTCAAGCTGGAACAACATCTCTTGGATGCCAAATGGGAATATCAAGATTG	641
Qy	582	CTAAGGCAATTTCTTCCGGCAGATCGTCTGTGCGTGCACTATTTGCATCAGAACAGAGTGTG	581
Db	522	CCAGCCGGGGGCGACTTTTATATCACTACGCGAGCGGCGACAGCTCAGTAGACGCGCAG	521
Qy	462	TCATTTGCCATTCATAGAAATGTTTGAACACAGACCAAGATGTATATCTCATGAGATATG	461
Db	402	AAGATCTGATCACTACGAGGAGATTTGATCATGTCACTCACTCAACCACTTCA	401
Qy	342	AGAGTGTGGGGCGCTGTGTGCGCATCAAGTCAATCCGGAAGACAAATCAAAATGATGAC	341
Db	282	GCTACAGATTCCTGTGAAGCCCTGTGGGCAAGGCACTTACGGGAAGGTGAAGAAAGGCGCGG	281
Qy	222	AGCCCTTATGAAGAGCAGCGGTGAAGCGGCACACACAGACACAACTGTGGGCAAC	221
Db	162	COACTCCCTCGGCGGACGAGCTACGCCGGCGCTGTGGGGAAGGGCTGATCAAGTGTGCCA	161

Db 1242 ATTGCTCAAGAAAGTCCCGCAAGAGAAATGACATGCCCCAGTCTCTCCACAGTGAACGG 1301
Qy 1187 CTGAGGATACCTCTTCTGCGCCCTGGCAAGACAGCTTAAAGCTTCCGAAAGGCAATTCTCA 1246
Db 1302 CTGATGACACCTGCCCATTCGCTTGGCAAGACAACTCAAGCTGCAAAAGGCAATTCTCA 1361
Qy 1247 AGAAAGAGCTCTTAACTCTGTCAGGGAGGTACAGAGAGACCTTCAGAACTCAAGACCG 1306
Db 1362 AGAAGAGAGTGTCAAGCTCTGTCAGAGGGGTACAGAGAGACCTTCGAGCTCAGCCCA 1421
Qy 1307 TGCTGTATCTTCCAGGGAGAGCTGTCTCTGCTGATCCCTCTCTCCCAAGAAAGCAATCC 1366
Db 1422 TCCCTGAGAGCCAGGGGAGGCTGCC-----CGCTCTCCCAAGAGGCAATTC 1472
Qy 1367 TTAAAGATCTCAAGAGGCAATCTGGTACTACTCTCTCCAGAGCCAGGAGTGTG 1426
Db 1473 TCAGAGAGCCCGCAAGGCGAGGTCTGGCTACTACTCTCTCCAGGCGCAAGTATGT 1532
Qy 1427 GGGAACTTTAAGACGCGCATGATGTGTTGTGAGTGGGAGCCCGTGGAGCAAGTCTC 1486
Db 1533 GGGAGCTCTTGAAGAGGCGACGATGTTGTGAGTGGGATCCCAAGAGCAAGCTTC 1592
Qy 1487 CACAGGCTTCAAGGAGCTCTCTCTCCAGGCAAGGCAATCTCAATCTCAAGTGTCT 1546
Db 1593 CGAAGCTTCAAGGAGCTCTCTCTCAATGCAAAAGGCAATCTCAATGCAAGTGTCT 1652
Qy 1547 CCGGACAGGCTTGAAGAGCACTACCCCTAGACCTTGGCTCCCTGGAGCAAGTGGCT 1606
Db 1653 CCGAGACGCTTGAAGTCTCGGAGCCCGCCACCACTTGGCTCCCTGGATGAATCTCGCTC 1712
Qy 1607 CTTCCCATCTGACGCGCGCCAGCGCCCTCAAGGAGCTGTGAGTGAAGACAGCATCC 1666
Db 1713 CACTCGCGCCCTGAGGCGCGGCGAGCGGACCTCAAGGAGCTGTGAGGAGAGACAGCATCC 1772
Qy 1667 TGCTCTCGAGTCTTTGACCAATTTGACCTTGCTGAAGCTTCTCCGAAACCCCACTGA 1726
Db 1773 TGCTCTCGAGTCTTTGACCAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1832
Qy 1727 GGGGCTGTGTCTGTGAGCAACTGAGGGGCTTGAAGCAAGCTTCCCTCAAG----- 1780
Db 1833 GGGGCTGTGTCTGTGAGCAACTGAGGGGCTTGAAGCAAGCTTCCCTCAAGGGGCTGTG 1892
Qy 1781 -----GTCTGAACGATGTGTGAGCAAGTCTTGGGGATGATGCTTTCTCTGACAG 1834
Db 1893 GAACTGCTGAGGCGCTGAGGAGGAGATCTTTGGGGAGAGAGCTTTTCTCTGACAG 1952
Qy 1835 ACTGCCAAGAGGTGATGACGCTTACAGACAAAGCTTGAAGTGTCTCAAGCTCAGC 1893
Db 1953 ACTGCCAAGAGGTGATGACGCTTACAGACAAAGCTTGAAGTGTCTCAAGCTCAGC 2011

RESULT 15
US-10-343-514-13
; Sequence 13, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBEYRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DEUCA/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13
; LENGTH: 1186
; TYPE: DNA
; ORGANISM: RAT
US-10-343-514-13

Query Match 56.6%; Score 1071.4; DB 18; Length 1186;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1114; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 70 GCCCGGCGCTGGGGAGCGGCTCATCAAGTCGCTTAACTGTGATGAAGAGCAGCG 129
Db 1 GCCCGGCGCTGGGGAGCGGCTCATCAAGTCGCTTAACTGTGATGAAGAGCAGCG 60

Qy 130 GTGAAGCGGACCATCAACAACAACCTGCGGACCGCTCAAGTTCCTGAGACGCTG 189
Db 61 GTGAAGCGGACCATCAACAACAACCTGAGGACCGCTCAAGTTCCTGAGACGCTG 120

Qy 190 GGCAGAGGCACTTACGGGAGGTGAAGAGGACGAGAGAGCTGGGCGCTGTGGTGC 249
Db 121 GGCAGAGGCACTTACGGGAGGTGAAGAGGACGAGAGAGCTGGGAGCGCTGGTGC 180

Qy 250 ATCAAGTTCATCAGAAAGAAACAATAATGACAGATGTGCTGCAATACAGAG 309
Db 181 ATCAAGTTCATCAGAAAGAAACAATAATGACAGATGTGCTGCAATACAGAG 240

Qy 310 GAGATTGAGATCATGTCTTCACTCAACAACCCCAATCATTTGCAATGAAAGTGT 369
Db 241 GAGATTGAGATCATGTCTTCACTCAACAACCCCAATCATTTGCAATGAAAGTGT 300

Qy 370 GAGATTGAGATCATGTGTGATGTCAATGAGATGTGCAAGCGAGCGATCTGTATG 429
Db 301 GAGATTGAGATCATGTGTGATGTCAATGAGATGTGCAAGCGAGCGATCTGTATG 360

Qy 430 TACATCAAGTGAAGCGGACCGCTGATGAGAGCGGAGCGGACGATTTTCCAGACATC 489
Db 361 TACATCAAGTGAAGCGGACCGCTGATGAGAGCGGAGCGGACGATTTTCCAGACATC 420

Qy 490 GTGTCTGCGCTGCACTACTGTGCAACAGAGGAGATGCTTCAACGAGATCTCAAGCTG 549
Db 421 GTGTCTGCGCTGCACTACTGTGCAACAGAGGAGATGCTTCAACGAGATCTCAAGCTG 480

Qy 550 AACATCTCTTGAATGCGCAATGAAACATCAAGATGCTGACTTTGGCTCTCAAGCTG 609
Db 481 AACATCTCTTGAATGCGCAATGAAACATCAAGATGCTGACTTTGGCTCTCAAGCTG 540

Qy 610 TACCACAAGGCAAGTCTTCAACAGCTTGTGTGAGAGCCCTCTTACGCTCGCTGAG 669
Db 541 TACCACAAGGCAAGTCTTCAACAGCTTGTGTGAGAGCCCTCTTACGCTCGCTGAG 600

Qy 670 ATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGCTCTGGGCGTCTC 729
Db 601 ATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGCTCTGGGCGTCTC 660

Qy 730 CTGTACATCTGTGTGATGAGCAACATGCGCTTGAACGGGAGAGATCATAAACAATG 789
Db 661 CTGTACATCTGTGTGATGAGCAACATGCGCTTGAACGGGAGAGATCATAAACAATG 720

Qy 790 AAGCAATCAAGTAAAGGGGCTTACCGTGAAGCGGCCCAAGCGCTGCATGCTGTGGCTG 849
Db 721 AAGCAATCAAGTAAAGGGGCTTACCGTGAAGCGGCCCAAGCGCTGCATGCTGTGGCTG 780

Qy 850 ATCCGCTGCTGTAAATGATGAACCCACCGCTGGGAGCAACAGAGAGATGATGAGCT 909
Db 781 ATCCGCTGCTGTAAATGATGAACCCACCGCTGGGAGCAACAGAGAGATGATGAGCT 840

Qy 910 CATTGCTGCTCAACTGGGCTTAAACAACCGGAGTGGGAGAACAGAGAGCTGCTGTAG 969
Db 841 CATTGCTGCTCAACTGGGCTTAAACAACCGGAGTGGGAGAACAGAGAGCTGCTGTAG 900

Qy 970 GGTGGGACCTTATGCTGATCTTTGGCGGCGCTTCAATGAGAGCTGATACGTGCTCC 1029
Db 901 GGTGGGACCTTATGCTGATCTTTGGCGGCGCTTCAATGAGAGCTGATACGTGCTCC 960

QY 1030 TCGGCGCCCTCTGGAATGAGCCAGGTGTGCACTTCTCAAGCAGACGTGGC 1089
 Db 961 TCCGCCCCCTCTGGAATGAGCCAAAGTGTGTAGCTTCTCAAGCAGACGTGGC 1020
 QY 1090 GAGGTGGAAGCACTGTACTGTGGCTGAGCGGCAACATTTCTTAAGAGTCCGGAAG 1149
 Db 1021 GAGGTGGAAGCAAGGACCGGGCTGAGCGGCAACATTTCTTAAGAGTCCGGAAG 1080
 QY 1150 GAGATGACATGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTTTCTGGCCCT 1209
 Db 1081 GAGATGACATGCTCAAGCTCTGAGACTGTGCAAGATGACCCAGTTGAAGATCTTCTCTGGCCCT 1140
 QY 1210 GGCAGAGCAGCCCTTAAGCTTCCGAAAGGCAATTCTCAAGAAAAAG 1254
 Db 1141 GGCAGAGCAGCCCTCAAGCTTCCGAAAGGTATCTCAAGAAAAAG 1185

Search completed: May 16, 2005, 09:29:19
 Job time : 1084.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:12:18 ; Search time 64.9672 Seconds
(without alignments)

1994.176 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGTGKVKAKRE.....VNPTRRATLEDVASHMWNVNW 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342	100.0	631	2 Q9DBV0	Q9DBV0 m mus muscu
2	1342	100.0	631	2 Q8CIC0	Q8CIC0 mus musculu
3	1338	99.0	639	2 Q80ZM3	Q80ZM3 mus musculu
4	1338	99.0	639	2 Q8BZN4	Q8BZN4 mus musculu
5	1313	97.8	630	2 Q66HE5	Q66HE5 rattus norv
6	1293	96.3	628	2 Q9H093	Q9H093 homo sapien
7	1128.5	84.1	661	1 ARK5 HUMAN	Q60285 homo sapien
8	1120.5	83.5	658	2 Q64IK5	Q64IK5 mus musculu
9	1025	76.4	575	2 Q616D6	Q616D6 mus musculu
10	824.5	61.4	383	2 Q7QBH4	Q7QBH4 anopheles g
11	822.5	61.3	1180	2 Q7KSS0	Q7KSS0 drosophila
12	822.5	61.3	1180	2 Q7YU19	Q7YU19 drosophila
13	822.5	61.3	1427	2 Q9VH05	Q9VH05 drosophila
14	813.5	60.6	1551	2 Q65ZH4	Q65ZH4 caenorhabdi
15	813.5	60.6	1592	2 Q65ZH3	Q65ZH3 caenorhabdi
16	665.5	49.6	752	1 MKR4 HUMAN	Q96134 homo sapien
17	665.5	49.6	752	2 Q8CIP4	Q8CIP4 mus musculu
18	660.5	49.2	752	2 Q8NG37	Q8NG37 homo sapien
19	652.5	48.6	725	2 Q804T2	Q804T2 xenopus lae
20	651.5	48.5	729	2 Q9JKE4	Q9JKE4 mus musculu
21	651.5	48.5	744	2 Q9JKE5	Q9JKE5 mus musculu
22	650	48.4	712	2 Q7QBH3	Q7QBH3 anopheles g
23	647.5	48.2	691	2 Q96RG0	Q96RG0 homo sapien
24	647.5	48.2	719	2 Q96RG0	Q96RG0 homo sapien
25	647.5	48.2	722	2 Q08679	Q08679 rattus norv
26	647.5	48.2	722	2 Q6PRD4	Q6PRD4 mus musculu
27	647.5	48.2	745	2 Q15524	Q15524 homo sapien
28	647.5	48.2	755	2 Q7KZ17	Q7KZ17 homo sapien
29	647.5	48.2	778	2 Q96HB3	Q96HB3 homo sapien
30	647.5	48.2	797	2 Q8VHFO	Q8VHFO rattus norv
31	647.5	48.2	888	2 Q8BR95	Q8BR95 mus musculu

32	644	48.0	512	2 P92958	P92958 arabidopsis
33	644	48.0	512	2 P92968	P92968 arabidopsis
34	643.5	48.0	722	2 Q80ZM0	Q80ZM0 brachydanio
35	642.5	47.9	792	2 Q61WT7	Q61WT7 xenopus lae
36	642.5	47.9	793	2 Q08678	Q08678 rattus norv
37	642.5	47.9	795	2 Q8VHJ5	Q8VHJ5 mus musculu
38	641.5	47.8	776	2 Q7ZYL7	Q7ZYL7 xenopus lae
39	641.5	47.8	785	2 Q8QGV3	Q8QGV3 xenopus lae
40	641	47.8	504	2 P93113	P93113 cucumis sat
41	640.5	47.7	462	2 Q75271	Q75271 homo sapien
42	640.5	47.7	542	2 Q6V8Y5	Q6V8Y5 physcomitre
43	639.5	47.7	780	2 Q804T1	Q804T1 xenopus lae
44	638.5	47.6	1398	2 Q77268	Q77268 drosophila
45	638.5	47.6	1398	2 Q9W532	Q9W532 drosophila

ALIGNMENTS

RESULT 1
Q9DBV0 PRELIMINARY; PRT; 631 AA.
ID Q9DBV0;
AC Q9DBV0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,
DE clone:1200013B22 product:weakly similar to PROBABLE SERINE/THREONINE-
DE PROTEIN KINASE KIAA0537 (EC 2.7.1.-) (Mus musculus adult male cecum
DE cDNA, RIKEN full-length enriched library, clone:9130215K18
DE KIAA0537).
DE Product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
DE KIAA0537).
GN Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RN [2]
RP STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RA The FANTOM Consortium.
RT "The RIKEN Genome Exploration Research Group Phase I & II Team:
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN [4]
RP STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RA MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RT Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RT Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subcloning of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RP STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.",
 RL Genome Res. 10:11757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito K., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cecum;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirouane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohata N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK004737; BAB23518.1; -
 DR EMBL; AK033672; BAC28421.1; -
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR00719; Prot. kinase.
 DR InterPro; IPR002290; Ser. Thr. kinase.
 DR InterPro; IPR008271; Ser. Thr. kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot. Kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 631 AA; 69806 MW; D51C042DE6CC174B CRC64;

Query Match 100.0%; Score 1342; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred. No. 2,6e-97;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGVKVKARESSGRLVAIKSRDKIKDEODLHIREIEMSLNHPH 60
 DB 57 YEFLETLGKGTGVKVKARESSGRLVAIKSRDKIKDEODLHIREIEMSLNHPH 116
 QY 61 IAIHEVFENSSKIVIMEYASRDLYDYISERPRISERDARHFFROIYSALHYCHONGIV 120
 DB 117 IAIHEVFENSSKIVIMEYASRDLYDYISERPRISERDARHFFROIYSALHYCHONGIV 176

QY 121 HRDLKLENTLLDANGNIKIADPGLSNLYHKGKFLQFPCCSPLYASPEIYNGKPYGPEVD 180
 DB 177 HRDLKLENTLLDANGNIKIADPGLSNLYHKGKFLQFPCCSPLYASPEIYNGKPYGPEVD 236
 QY 181 SMSLGLVLYLYVHGTPFPQGDHRTLVKQISNGAYEPPKPSDAGLIRWLWVNPERRA 240
 DB 237 SMSLGLVLYLYVHGTPFPQGDHRTLVKQISNGAYEPPKPSDAGLIRWLWVNPERRA 296
 QY 241 TLEDVASHWVWVW 253
 DB 297 TLEDVASHWVWVW 309
 RESULT 2
 ID 08C1C0 PRELIMINARY; PRT; 631 AA.
 AC 08C1C0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 1200013B22Rik protein.
 GN Name=1200013B22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen R.D., Mullany S.J.,
 RA Altshul S.F., Zeeberg B., Bueltow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Mususina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidl T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Czech II; TISSUE=Mammary tumor;
 RA Strausberg R.L.
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC033302; AAH33302.1; -
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR00719; Prot. kinase.
 DR InterPro; IPR002290; Ser. Thr. kinase.
 DR InterPro; IPR008271; Ser. Thr. kinase.
 DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot. Kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 631 AA; 69745 MW; C6C9042247D9174F CRC64;

Query Match 100.0%; Score 1342; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 2, 6e-97;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 60
DB YEFLETLGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 116
QY 61 IAIHEVENSCKIVIMEYASRGDLYDISERPRLSERDARHFRQIVSAHYCHONGIV 120
DB IAIHEVENSCKIVIMEYASRGDLYDISERPRLSERDARHFRQIVSAHYCHONGIV 176
QY 121 HRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGKPYVGEVD 180
DB HRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGKPYVGEVD 236
QY 177 HRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGKPYVGEVD 240
DB 181 SMSLGVLLYLIVHGTMPFDGQDHKTIVKQISNGAYRBPSPDAGLIRWLLWNPTRRA 240
QY 237 SMSLGVLLYLIVHGTMPFDGQDHKTIVKQISNGAYRBPSPDAGLIRWLLWNPTRRA 296
DB 237 SMSLGVLLYLIVHGTMPFDGQDHKTIVKQISNGAYRBPSPDAGLIRWLLWNPTRRA 296
QY 241 TLEDVASHWVWVW 253
DB 297 TLEDVASHWVWVW 309

RESULT 3

ID 0802W3 PRELIMINARY; PRT; 639 AA.

AC 0802W3;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE 1200013B22R1k protein.
GN Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hults S.W.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tohbiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.J., Gay L.J., Hults S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Krzyzanski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.B.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX Straubeberg R.;
RA Submitted (PEB-2003) to the EMBL/GenBank/DBJ databases.
CC -/- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; BC046833; AA046833.1; -;
DR HSP; P31751; IGZK.

DR MED; MG1.1921387; 1200013B22R1k.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004688; P:Protein amino acid phosphorylation; IEA.

DR InterPro; IPR001109; Kinase like.
DR InterPro; IPR000719; Prot. kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR Pfam; PF00069; Kinase; 1.
DR Prodom; PD000001; Prot. kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 639 AA; 70675 MW; FB9C40228F53872C CRC64;

Query Match 99.0%; Score 1328; DB 2; Length 639;
Best Local Similarity 96.9%; Pred. No. 3, 3e-96;
Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 60
DB YEFLETLGKGTGKVKKARSSGRLVAKISIRDKIKDEODLHIREIIMSLNPHI 116
QY 61 IAIHE-----VFENSCKIVIMEYASRGDLYDISERPRLSERDARHFRQIVSAH 112
DB IAIHEVGRSLVIVFENSSCKIVIMEYASRGDLYDISERPRLSERDARHFRQIVSAH 176
QY 117 IAIHEVGRSLVIVFENSSCKIVIMEYASRGDLYDISERPRLSERDARHFRQIVSAH 176
DB 117 YCHONGIVHRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 172
QY 173 YCHONGIVHRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 236
DB 177 YCHONGIVHRDLKLENILLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIYVNGK 236
QY 173 PYGPEVDSWSLGVLLYLIVHGTMPFDGQDHKTIVKQISNGAYRBPSPDAGLIRWLL 232
DB 237 PYGPEVDSWSLGVLLYLIVHGTMPFDGQDHKTIVKQISNGAYRBPSPDAGLIRWLL 296
QY 233 MNVPTRRATLEDVASHWVWVW 253
DB 297 MNVPTRRATLEDVASHWVWVW 317

RESULT 4

ID 08B2N4 PRELIMINARY; PRT; 639 AA.

AC 08B2N4;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:933015AN24 product:weakly similar to PROBABLE
DE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.1.-).
GN Name=1200013B22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:119-44(1999).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RA The PANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirokane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Morimatsu M., Hayashizaki Y.;
 RT Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- Similarity: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK034082; BAC26575.1; -.
 DR HSSP; P31751; IGZK.
 DR MGD; MGI:1921387; 1200013B22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR002230; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 639 AA; 70632 MW; F898BEE915C95F85D CRC64;

Query Match 99.0%; Score 1328; DB 2; Length 639;
 Best Local Similarity 96.9%; Pred. No. 3.3e-96;
 Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YEFLETLGKGTGKVKARESSGRLVAIKSRDKIKDEBDLHIREIEMSLNHPH 60
 DB 57 YEFLETLGKGTGKVKARESSGRLVAIKSRDKIKDEBDLHIREIEMSLNHPH 116

QY 61 IAIIHE-----VFENSSKIYIMVYASRGDIYDISSERPRISERDARHFFQIYSALH 112
 DB 117 IAIIHEVGRSRLVVFENSSKIYIMVYASRGDIYDISSERPRISERDARHFFQIYSALH 176
 QY 113 YCHQNGIVHRDLKLENIILDANGNIIKIADFGSLNLYHKGFLOTFCGSPLYASPELVNGK 172
 DB 177 YCHQNGIVHRDLKLENIILDANGNIIKIADFGSLNLYHKGFLOTFCGSPLYASPELVNGK 236
 QY 173 PYVGPEVDSWSLGLVLLIIVHGTMPDGDHKTLYVQISNGAYREPPKPSDACGLIRWL 232
 DB 237 PYVGPEVDSWSLGLVLLIIVHGTMPDGDHKTLYVQISNGAYREPPKPSDACGLIRWL 296
 QY 233 MNVPTRRATLEDVASHWVNW 253
 DB 297 MNVPTRRATLEDVASHWVNW 317

RESULT 5
 Q66HE5 PRELIMINARY; PRT; 630 AA.
 ID Q66HE5
 AC Q66HE5;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang Y., Hsieh F.,
 RA Diatchenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.R., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Munz D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Director MGC Project;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC081899; AAH81899.1;
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR002230; Prot_kinase.
 DR InterPro; IPR002230; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TYKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein.

SQ SEQUENCE 630 AA; 69952 MW; 1626427ABCD5F66E CRC64;
Query Match 97.8%; Score 1313; DB 2; Length 630;
Best Local Similarity 98.0%; Pred. No. 5e-95;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRDKIKDEODLHIREIEMSLNPHI 60
DB 57 YEFLETLGKGTGKVKKARSSGRLVAIKSIRDKIKDEODLHIREIEMSLNPHI 116
QY 61 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 120
DB 117 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNKIKADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
DB 177 HRDLKLENTLLDANGNKIKADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 236
QY 181 SMSLGVLLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPSPDACGLIRMLVMNPTERRA 240
DB 237 SMSLGVLLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPSPDACGLIRMLVMNPTERRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309
RESULT 6
ID Q9H093 PRELIMINARY; PRT; 628 AA.
AC Q9H093;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp434J037 (Hypothetical protein SNARK)
GN (Hypothetical protein DKFZp434J037; Synonymus=SNARK;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RG The German cDNA Consortium;
RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fodor G., Han M., Wiemann S.,
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Yoshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Krzywinski M.I., Skaleka U., Smillie D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
RN [3]

RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Isegaki T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hito Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- Similarity: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AL136891; CAB6825.1; -!
DR EMBL; BC017306; AAH17306.1; -!
DR EMBL; AK074830; BAC11234.1; -!
DR HSSP; P1751; 1GZK
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot. kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot. kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 628 AA; 69611 MW; F76FB1BF9AF4C67 CRC64;
Query Match 96.3%; Score 1293; DB 2; Length 628;
Best Local Similarity 95.7%; Pred. No. 1.9e-93;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRDKIKDEODLHIREIEMSLNPHI 60
DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRDKIKDEODLHIREIEMSLNPHI 112
QY 61 IAIHEVENSCKIVIMEVYASRGDLVDYISERPLSRDARHFRQIVSAHYCHONGIV 120
DB 113 IAIHEVENSCKIVIMEVYASRGDLVDYISERQLSREARHFRQIVSAHYCHONGIV 172
QY 121 HRDLKLENTLLDANGNKIKADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
DB 173 HRDLKLENTLLDANGNKIKADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 232
QY 181 SMSLGVLLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPSPDACGLIRMLVMNPTERRA 240
DB 233 SMSLGVLLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPSPDACGLIRMLVMNPTERRA 292
QY 241 TLEDVASHMWVNM 253
DB 293 TLEDVASHMWVNM 305
RESULT 7
ARKS HUMAN
ID ARKS HUMAN STANDARD; PRT; 661 AA.
AC 060265;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE AMPK-related protein kinase 5 (EC 2.7.1.37).
GN Name=ARK5; Synonymus=KIAA0537;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
RC TISSUE=Brain.
RA MEDLINE=9820545; PubMed=9628581;
RA Nagae T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [2]
RP CHARACTERIZATION, AND MUTAGENESIS OF SER-600.
RX MEDLINE=22393479; PubMed=12409306; DOI=10.1074/jbc.M206025200;
RA Suzuki A., Kusakai G.-I., Kishimoto A., Lu J., Ogura T., Tavin M.F.,
RA Esauli H.;
RT "Identification of a novel protein kinase mediating Akt survival
RT signaling to the ATM protein.";
RL J. Biol. Chem. 278:48-53(2003).
CC -1- FUNCTION: Involved in tolerance to glucose starvation.
CC Phosphorylates ATM.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: Activated by PKA/AKT1 during glucose
CC starvation.
CC -1- TISSUE SPECIFICITY: Expressed at high levels in heart and brain,
CC and at lower levels in skeletal muscle, kidney, ovary, placenta,
CC lung and liver.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011109; BAA25463.1; -.
DR HSSP; P31751; IGZK.
DR MIM; 608130; -.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; P080108; Phosphorylation; Polymorphism;
DR Serine/threonine-protein kinase; Transferase.
FT DOMAIN 55 306 Protein kinase.
FT NP BIND 61 69 ATP (By similarity).
FT ACT_SITE 178 178 Proton acceptor (By similarity).
FT BINDING 84 84 ATP (By similarity).
FT MOD_RES 600 600 Phosphoserine (by PKA/AKT1).
FT VARIANT 543 543 P -> R (in dbSNP:3741883).
FT /FTID=VAR_017246.
FT MUTAGEN 600 600 S->A: No phosphorylation.
SQ SEQUENCE 661 AA; 74304 MW; 806F37D52CA4718F CRC64;
Query Match 84.1%; Score 1128.5; DB 1; Length 661;
Best Local Similarity 82.3%; Pred. No. 1.8e-80; Indels 1; Gaps 1;
Matches 209; Conservative 24; Mismatches 20;
QY 1 YEFLETLGKGTGKVKARBS-SGRVLAIKSRKDKIKOEQDLHRIREIEMSLNHP 59
DB 55 YELQETLGKGTGKVKARERFSGRVAIKSRKDKIKOEQDWHIRREIEMSLNHP 114
QY 60 IAIHVEFENSKIVTVMEYASRGDLYDISRPRLSERDARHPFQIVSAVHYCHQNGI 119
DB 115 IISIEVEFENKDKIVIMEYASRGELDYDISERRRISERTRHFPQIVSAVHYCHQNGV 174
QY 120 VHRDLKLENIILDANGNINIKIADFGSLNLYHKGFLOTFCGSPLYASPEIVNGRPYGPV 179

DB 175 VHRDLKLENIILDANGNINIKIADFGSLNLYXQKFLOTFCGSPLYASPEIVNGRPYGPV 234
QY 180 DSWSLGVLIIYLVHGTMPDPGDDHKTLYKQISNGAYREPPKPSDACGLIRMLMNPTR 239
DB 235 DSWALGVLLIYLVYGTMPDPGFDHKLIRQISGSEYREPTQPSDARGLIRMLMNPDR 294
QY 240 ATLEDVASHMWNV 253
DB 295 ATLEDIANHWNV 308
RESULT 8
ID 0641K5 PRELIMINARY; PRT; 658 AA.
AC 0641K5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE RIKEN cDNA B230104P22.
GN Name=B230104P22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schenck A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082328; AAH82328.1; -.
SQ SEQUENCE 658 AA; 73661 MW; E7176F13B75B889F CRC64;
Query Match 83.5%; Score 1120.5; DB 2; Length 658;
Best Local Similarity 81.5%; Pred. No. 7.9e-80;
Matches 207; Conservative 25; Mismatches 21; Indels 1; Gaps 1;
QY 1 YEFLETLGKGTGKVKARBS-SGRVLAIKSRKDKIKOEQDLHRIREIEMSLNHP 59
DB 56 YELQETLGKGTGKVKARERFSGRVAIKSRKDKIKELDMWHIRREIEMSLNHP 115
QY 60 IAIHVEFENSKIVTVMEYASRGDLYDISRPRLSERDARHPFQIVSAVHYCHQNGI 119
DB 116 IISIEVEFENKDKIVIMEYASRGELDYDISERRRISERTRHFPQIVSAVHYCHQNGV 175
QY 120 VHRDLKLENIILDANGNINIKIADFGSLNLYHKGFLOTFCGSPLYASPEIVNGRPYGPV 179
DB 176 VHRDLKLENIILDANGNINIKIADFGSLNLYXQKFLOTFCGSPLYASPEIVNGRPYGPV 235
QY 180 DSWSLGVLIIYLVHGTMPDPGDDHKTLYKQISNGAYREPPKPSDACGLIRMLMNPTR 239

Db 236 DSWALGVLLYTLTYGMPDPGFDHKLIRQISGSEYREPTQPSDAGLIRMLMNVNDR 295
QY 240 ATLEDVASHMWNV 253
Db 296 ATIEDIANHWNV 309

RESULT 9
ID 0616D6 PRELIMINARY; PRT; 575 AA.
AC 0616D6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
DE MKIAA0537 protein (Fragment).
GN Name=mkIAA0537;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Iamato S., Koseki H., Hiraoka S.,
Saga Y., Hagiwara T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIA Gene:
RT III. The Complete Nucleotide Sequences of 500 Mouse KIA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Sire-fractionated Libraries.";
RL DNA Res. 10:167-180(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Nagase T., Ohara O., Koga H.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB182364; BAD2395.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 575 AA; 64949 MW; 3CA0796FBA199D2B CRC64;

Query Match 76.4%; Score 1025; DB 2; Length 575;
Best Local Similarity 82.3%; Pred. No. 2.3e-72;
Matches 186; Conservative 23; Mismatches 17; Indels 0; Gaps 0;

QY 28 IKSIRKDKIKEDODLHIREIEIMSSLNHPHIAIHEVENSKIYIVMEYASRGDLYD 87
Db 1 IKSIRKDKIKEDODLHIREIEIMSSLNHPHIAIHEVENSKIYIVMEYASRGDLYD 60
QY 88 YISERPRISERDARHFPROIVSALAYCHONGIYVHDKLENIILLDANGNIKIADFGLSNL 147
Db 61 YISERPRISERDARHFPROIVSALAYCHONGIYVHDKLENIILLDANGNIKIADFGLSNL 120
QY 148 YHKGKFLQTFCCGSLVASPEIVNGKPYGVDSVMSGLVLLYLVHGTMPEPDQDKTIV 207
Db 121 YQDKKFLQTFCCGSLVASPEIVNGKPYGVDSVMSGLVLLYLVHGTMPEPDQDKKLL 180
QY 208 KOISNGAYREPPKPSDAGLIRMLMNVNPTRRATLEDVASHMWNV 253

Db 181 KOISGSEYREPTQPSDAGLIRMLMNVNPTRRATLEDIANHWNV 226
RESULT 10
ID 070BH4 PRELIMINARY; PRT; 383 AA.
AC 070BH4
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)
DE Eb1p238 (Fragment).
GN Name=Eb1p238; ORFNames=ENSGANG00000012297;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008879; EAA08421.1; -.
DR HSSP; P31751; IGZK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
FT NON_TER 383
SQ SEQUENCE 383 AA; 43133 MW; 294732D45CFC1023 CRC64;

Query Match 61.4%; Score 824.5; DB 2; Length 383;
Best Local Similarity 61.3%; Pred. No. 9e-57;
Matches 155; Conservative 44; Mismatches 53; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVKKA-RESSGRVIAIKSRDKIKEDODLHIREIEIMSSLNHPH 59
Db 44 FDIKTLGGGTGKQVGLINKETGCEVALKITKSKIEFEDLIRRRVQIMSSVQHRN 103
QY 60 IIAIHEVENSKIYIVMEYASRGDLYDYSERPRISERDARHFPROIVSALAYCHONGI 119
Db 104 IIAIHEVENSKIYIVMEYASRGDLYDYSERPRISERDARHFPROIVSALAYCHONGI 163
QY 120 VHRDLKLENIILLDANGNIKIADFGLSNLVHKGKFLQTFCCGSLVASPEIVNGKPYGVDS 179
Db 164 CHRDLKLENIILLDANGNIKIADFGLSNLVHKGKFLQTFCCGSLVASPEIVNGKPYGVDS 223
QY 180 DMSVGLVLLYLVHGTMPEPDQDKTIVKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 239
Db 224 DMSVGLVLLYLVHGTMPEPDQDKTIVKQISNGAYREPPKPSDAGLIRMLMNVNPTRR 283
QY 240 ATLEDVASHMWNV 252
Db 284 ANIEQICNHNWNV 296

RESULT 11
ID 07KS60 PRELIMINARY; PRT; 1180 AA.

DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
LN 05937D.
GN ORFNames=CG11870;
DN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez C.J., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Pargass V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,
RA Celniker S.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: B010033; AAQ2502.1; -.
DR HSSP: P31751; 1MRV.
DR Flybase: FBgn0037804; CG11870.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR001109; Kinase-like.
DR InterPro: IPR000719; Prot_Kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1180 AA; 127684 MW; 2CD790A4F3BDC50C CRC64;

Query Match 61.3%; Score 822.5; DB 2; Length 1180;
Best Local Similarity 61.7%; Pred. No. 4.8e-56;
Matches 156; Conservative 44; Mismatches 52; Indels 1; Gaps 1;

QY 1 YEFLETLGKGYCKYKKA-RESSGRVATKSTKDKIXOBODLHTRREIEMSSLNHPH 59
DB 70 FIIKKLGGGYGKQVGLINKETGQGVATKTKCKIEADLVRIIRREVQIMSVHHP 129
QY 60 IIAIHEVFENSSKIVIVMEYASRGDLYDISSRPRLSEDRARHFPQIVSALHYCHONGI 119
DB 130 IHHIVFENREKVMVMEFPAAGELIYDLSEKVKLTSEBARKIFQVATVYCKKHKI 189
QY 120 VHRDLKLENTLLDANGNIKADFGSLNLYHKRGLQTFGSPDYASPEIYNGKPYGPEV 179
DB 190 CHRDLKLENTLLDEKNAKIDFGSLNVDQRLTGFCGSPDYASPEIYEGPYGPEV 249
QY 180 DMSVGLVLLIYLVHGMPDPGDHKLTVQISNGAYRPEPKSDACGLIRMLMNVNPTKR 239
DB 250 DMSVGLVLLIYLVHGMPDPGDFGKRLVQISGDDYEPKPSRASTLIRMLTVCPRRK 309
QY 240 ATLEDAVSHMWNV 252
DB 310 ASIEQICSHMWN 322

RESULT 13
ID 09VH05 PRELIMINARY; PRT; 1427 AA.
AC 09VH05; 09VH04;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)

DE CG11870-PA (CG11870-pb).
GN ORFNames=CG11870;
DN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
[1]
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abiri J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostel A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kasko P.B., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Li X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mikhlov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Benman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.U., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN
RN
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN
RN SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB003686; AAP54517.3; -.
DR HSSP; O63450; 1A06.
DR Inactive; Q9VH05; -.
DR FlyBase; F8gn0037804; CG11870.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1427 AA; 153443 MW; 530928D0A057B5C8 CRC64;

Query Match 61.3%; Score 822.5; DB 2; Length 1427;
Best Local Similarity 61.7%; Pred. No. 6e-56;
Matches 156; Conservative 44; Mismatches 52; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGYKVKKA-RESSGRVIAIKSIRKDKIKDEODLHIREIEIMSSINHPH 59
DB 70 FPIIKLGGGTGKQVLGINKETGEVATIKTKKITEADLVIRIRREVQIMSVHNPN 129
QY 60 IAIHVEFENSSKIYIVMEYASRGDIYDYSERPRLSERDARHPFQIYSALHYCHQNGI 119
DB 130 IHIHVEFENRREBMVLMVEFAAGELDYLSERKVLTEEARIRIFQVATAVYCHKHKI 189
QY 120 VHRDLKLENIILLDANNIKIADFGLSNLVHKGFLOTFGSSPLVYASPELVNGEPYGPV 179
DB 190 CHRDLKLENIILLDEKNAKIALDGLSNVDPDQRLTGFGSSPLVYASPELVNGEPYGPV 249
QY 180 DMSLGLVLLIYLVHGTMPFGDQDKTLVQIISNGAYREPPKPSDAGLIRMLVMNPTPR 239
DB 250 DCMISGLVLLIYLVGSMFPDGSNFKRLVQIISGQDYERPKPSRASTLIRDMLTVCPRK 309
QY 240 ATLEDVASHMWVN 252
DB 310 ASIEQICSHMWVN 322

RESULT 14
Q65ZH4 PRELIMINARY; PRT; 1551 AA.
AC Q65ZH4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3.
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxId=6239;
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG MEDLINE=9069613; PubMed=9851916;
RX Wormbase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RT "The sequence of C. elegans cosmid B0496.";
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RN Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58749; AAU05597.1; -.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
SQ SEQUENCE 1551 AA; 176395 MW; 1F3BE62632A44C85 CRC64;

Query Match 60.6%; Score 813.5; DB 2; Length 1551;
Best Local Similarity 61.7%; Pred. No. 3.4e-55;
Matches 156; Conservative 33; Mismatches 63; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGYKVKARESS-GRVIAIKSIRKDKIKDEODLHIREIEIMSSINHPH 59
DB 75 FEITTKLGGGTGKSLAVDHFDEEVANVKKLKAISBKADLVIRIRIRIMSLNHPN 134
QY 60 IAIHVEFENSSKIYIVMEYASRGDIYDYSERPRLSERDARHPFQIYSALHYCHQNGI 119
DB 135 IIOIYVEFENKQKIIIVMEYSSGGEIYDVYVSCSGSLPEAEARIRIFQIITSAYLVYCHKHV 194
QY 120 VHRDLKLENIILLDANNNAKIALDFGLSNVPADKNLLTTFGSSPLVYASPELVNGEPYGPV 179
DB 195 AHRDLKLENIILLDNNNAKIALDFGLSNVPADKNLLTTFGSSPLVYASPELVNGEPYGPV 254
QY 180 DMSLGLVLLIYLVHGTMPFGDQDKTLVQIISNGAYREPPKPSDAGLIRMLVMNPTPR 239
DB 255 DCMISGLVLLIYLVGSMFPDGSNFKRLVQIISGQDYERPKPSRASTLIRDMLTVCPRK 314
QY 240 ATLEDVASHMWVN 252
DB 315 ATIFDIASHMWVN 327

RESULT 15

```
Q65ZH3
ID Q65ZH3 PRELIMINARY; PRT; 1592 AA.
AC Q65ZH3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B0496.3.
GN Name=B0496.3; ORFNames=B0496.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peioderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; Pubmed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Le T.T.;
RT "The sequence of C. elegans cosmid B0496.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; US8749; AAU05598.1; -.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Hypothetical protein.
SQ SEQUENCE 1592 AA; 180962 MW; E28D6FF934C43F76 CRC64;

Query Match 60.6%; Score 813.5; DB 2; Length 1592;
Best Local Similarity 61.7%; Pred. No. 3.5e-55;
Matches 156; Conservative 33; Mismatches 63; Indels 1; Gaps 1;

QY 1 YEFLETLGKGTGKVGKARSS-CRLVAIKSIRKDKIKDQDLHIREIEMSLNPH 59
DB 75 FEITTKLGSGTYGSLAYDHKEDREVAKLIRKSAIESKADLVIRIREIRIMSLNPHN 134
QY 60 IIAIHEVFENSXKIVIMVEYASRGDLVYISERPRLSERDARHPFQIVSALHYCHONGI 119
DB 135 IIQIYVEFENKDKIIVMEYSSGSELYDYVSRGSLPEAEKRRIFRQITSAVLXCHKHRY 194
QY 120 VHRDLKLENTILLDANGIKIADFGLSNLVHKGFLOTFCGSLTYASPEIYNGKPYGPEV 179
DB 195 AHRDLKLENTILLDQNNNAKIADFGLSNYFADKMLITTFCSPLTASPEIINGTPYKGPYV 254
QY 180 DSNSLGLVLYIVGTMPPGQDHKTLYKOISNGAYRREPXPSPDAGLIRMLLVNPNPTR 239
DB 255 DCMSLGILTYLTVGSMFQGRDPRNRVROIKRGATFEPEPTSTASMLIRMLRVNPNRR 314
QY 240 ATLEDVASHMWVN 252
```

```
DB 315 ATLEDVASHMWVN 327
```

```
Search completed: May 11, 2005, 14:24:01
Job time : 65.9672 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:24:11 ; Search time 52.0882 Seconds
(without alignments)
1622.537 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGTGKVKAKARE.....VNPTRRATLEDVASHMWVNW 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	14	US-10-355-975-11 Sequence 11, Appl
2	1325	98.7	251	16	US-10-343-514-103 Sequence 103, Appl
3	1313	97.8	630	16	US-10-343-514-41 Sequence 41, Appl
4	1296	96.6	251	16	US-10-343-514-50 Sequence 50, Appl
5	1293	96.3	628	9	US-09-963-159-2 Sequence 2, Appl
6	1293	96.3	628	15	US-10-423-543-44 Sequence 44, Appl
7	1293	96.3	672	16	US-10-618-941-77 Sequence 77, Appl
8	1286	95.8	594	15	US-10-311-034-112 Sequence 12, Appl
9	1276	95.1	251	16	US-10-343-514-101 Sequence 101, Appl
10	1273	94.9	640	16	US-10-322-281-23 Sequence 23, Appl
11	1230	91.7	616	16	US-10-322-281-26 Sequence 26, Appl
12	1128.5	84.1	661	9	US-09-780-949-2 Sequence 2, Appl
13	1128.5	84.1	661	9	US-09-780-949-6 Sequence 6, Appl

14	1128.5	84.1	661	14	US-10-354-358-82	Sequence 82, Appl
15	1111.5	82.8	252	16	US-10-343-514-102	Sequence 102, Appl
16	814	60.7	530	9	US-09-836-392-20	Sequence 20, Appl
17	785.5	58.5	246	10	US-09-898-8378-29	Sequence 29, Appl
18	665.5	49.6	508	15	US-10-016-248-24	Sequence 24, Appl
19	665.5	49.6	639	15	US-10-016-248-71	Sequence 71, Appl
20	665.5	49.6	639	15	US-10-016-248-72	Sequence 72, Appl
21	665.5	49.6	688	15	US-10-276-645-8	Sequence 8, Appl
22	665.5	49.6	703	15	US-10-016-248-70	Sequence 70, Appl
23	665.5	49.6	752	9	US-09-835-081-2	Sequence 2, Appl
24	665.5	49.6	752	15	US-10-258-106-16	Sequence 16, Appl
25	665.5	49.6	752	15	US-10-276-645-7	Sequence 7, Appl
26	665.5	49.6	769	15	US-10-363-616-403	Sequence 403, Appl
27	665.5	49.6	825	15	US-10-425-114-54516	Sequence 54516, A
28	661.5	49.3	752	16	US-10-618-941-79	Sequence 79, Appl
29	660.5	49.2	639	15	US-10-016-248-73	Sequence 73, Appl
30	660.5	49.2	660	15	US-10-276-645-6	Sequence 6, Appl
31	660.5	49.2	688	15	US-10-161-565-28	Sequence 28, Appl
32	660.5	49.2	688	15	US-10-161-565-29	Sequence 29, Appl
33	660.5	49.2	724	14	US-10-276-645-5	Sequence 5, Appl
34	651.5	48.5	744	9	US-09-835-081-4	Sequence 4, Appl
35	649.5	48.4	729	14	US-10-142-356-11	Sequence 11, Appl
36	649.5	48.4	729	14	US-10-195-101-33	Sequence 33, Appl
37	649.5	48.4	729	14	US-10-161-565-26	Sequence 26, Appl
38	649.5	48.4	744	9	US-09-919-585-3	Sequence 3, Appl
39	647.5	48.2	691	9	US-09-919-585-6	Sequence 6, Appl
40	647.5	48.2	722	15	US-10-274-194-4	Sequence 4, Appl
41	647.5	48.2	722	16	US-10-760-407-4	Sequence 4, Appl
42	647.5	48.2	724	9	US-09-919-585-9	Sequence 9, Appl
43	647.5	48.2	745	14	US-10-195-101-36	Sequence 36, Appl
44	647.5	48.2	745	14	US-10-161-565-24	Sequence 24, Appl
45	647.5	48.2	745	15	US-10-260-708-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-10-355-975-11
Sequence 11, Application US/10355975
Publication No. US20030162277A1
GENERAL INFORMATION:
APPLICANT: Immunex Corporation
APPLICANT: Bird, Timothy A.
APPLICANT: Virca, G. Duke
APPLICANT: Martin, Unja
TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
FILE REFERENCE: 2923-A
CURRENT APPLICATION NUMBER: US/10/355,975
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US/09/579,664B
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 631
TYPE: PRT
ORGANISM: Mus musculus
US-10-355-975-11

Query Match 100.0%; Score 1342; DB 14; Length 631;
Best Local Similarity 100.0%; Pred No. 1.2e-100;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YEFLETLGKGTGKVKAKARESSGRLVAIKSRDKIKDEODLHRRRIEIMSLNHPH 60
DB 57 YEFLETLGKGTGKVKAKARESSGRLVAIKSRDKIKDEODLHRRRIEIMSLNHPH 116
QY 61 IAIHEVFENSSKIVIMEASRDLDYIISERPRLSERDARHFRQIVSALHYCHONGIV 120
DB 117 IAIHEVFENSSKIVIMEASRDLDYIISERPRLSERDARHFRQIVSALHYCHONGIV 176

QY 121 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
| | | | |
DB 177 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
| | | | |
QY 181 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 240
| | | | |
DB 237 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 296
| | | | |
QY 241 TLEDVASHMWVW 253
| | | | |
DB 297 TLEDVASHMWVW 309
| | | | |
RESULT 2
US-10-343-514-103
; Sequence 103, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 103
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-514-103
Query Match 98.7%; Score 1325; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-99;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YEFLETGKGTGKVKARESSGRLVAIKSIRKDKIKDEQDLHIREIEIMSSLNHPHI 60
| | | | |
DB 1 YEFLETGKGTGKVKARESSGRLVAIKSIRKDKIKDEQDLHIREIEIMSSLNHPHI 60
| | | | |
QY 61 IAIHEVFENSSKIYVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
| | | | |
DB 61 IAIHEVFENSSKIYVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
| | | | |
QY 121 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
| | | | |
DB 121 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
| | | | |
QY 181 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 240
| | | | |
DB 181 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 240
| | | | |
QY 241 TLEDVASHMWV 251
| | | | |
DB 241 TLEDVASHMWV 251
| | | | |
RESULT 3
US-10-343-514-41
; Sequence 41, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
US-10-343-514-41

; APPLICANT: LEFEVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 41
; LENGTH: 630
; TYPE: PRT
; ORGANISM: RAT
US-10-343-514-41
Query Match 97.8%; Score 1313; DB 16; Length 630;
Best Local Similarity 98.0%; Pred. No. 2.7e-98;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 YEFLETGKGTGKVKARESSGRLVAIKSIRKDKIKDEQDLHIREIEIMSSLNHPHI 60
| | | | |
DB 57 YEFLETGKGTGKVKARESSGRLVAIKSIRKDKIKDEQDLHIREIEIMSSLNHPHI 116
| | | | |
QY 61 IAIHEVFENSSKIYVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
| | | | |
DB 117 IAIHEVFENSSKIYVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 176
| | | | |
QY 121 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
| | | | |
DB 177 HRDLKLENIILLDANGNIKADPGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
| | | | |
QY 181 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 240
| | | | |
DB 237 SMSLGLVLYLVHGTMPFDGQDHKTLYVKQISNGAYREPPKPSDACGLIRMLVMVNPTRRA 296
| | | | |
QY 241 TLEDVASHMWVW 253
| | | | |
DB 297 TLEDVASHMWVW 309
| | | | |
RESULT 4
US-10-343-514-50
; Sequence 50, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 50
; LENGTH: 251
; TYPE: PRT
; ORGANISM: RAT
US-10-343-514-50

Query Match Similarity	96.6%	Score 1296	DB 16	Length 251
Best Local Similarity	98.0%	Pred. No. 2.3e-97		
Matches	246	Conservative	3	Mismatches 2; Indels 0; Gaps 0
Qy	1	YEFLETLKGTGYKVKKARSSGRVAIAKSIKDKIKDEODLHTRREIEIMSSLNHPHI	60	
Db	1	YEFLETLKGTGYKVKKARSSGRVAIAKSIKDKIKDEODLHTRREIEIMSSLNHPHI	60	
Qy	61	IAIHVEFNSSKIVIVMEYASRGDLVDYISERPLRSERDARHFFQIVSALHYCHONGIV	120	
Db	61	IAIHVEFNSSKIVIVMEYASRGDLVDYISERPLRNERDARHFFQIVSALHYCHONGIV	120	
Qy	121	HRDLKLENLIDPANGNIKIADPGLSNLYHKKFLQTFCCSPLYASPELYNGKPYGPEYD	180	
Db	121	HRDLKLENLIDPANGNIKIADPGLSNLYHKKFLQTFCCSPLYASPELYNGKPYGPEYD	180	
Qy	181	SMSLGLVLLYIIIVHGMPPDGODHKLTVNQISGAAREPKPSDAGCLRILMLNVPTRRA	240	
Db	181	SMSLGLVLLYIIIVHGMPPDGODHKLTVNQISGAAREPKPSDAGCLRILMLNVPTRRA	240	
Qy	241	TLEDVASHMVV	251	
Db	241	TLEDVASHMVV	251	

```

RESULT 5
US-09-963-159-2
Sequence 2, Application US/09963159
Patent No. US20020077312A1
GENERAL INFORMATION:
APPLICANT: Curtiss, Rory A.J.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
FILE REFERENCE: 10147-5001
CURRENT APPLICATION NUMBER: US/09/963,159
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/234,922
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 628
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-159-2

```

Query Match	96.3%	Score 1293	DB 9	Length 628
Best Local Similarity	95.7%	Pred. No. 1	le-96	
Matches	242	Conservative	6	Mismatches 5, Indels 0, Gaps 0

QY	1	YEFLETLGKGYGKAYKARSSGRLVAIKSIRKDKIKOEODILHIREIETMSSLNPHI	60
DB	53	YEFLETLGKGYGKAYKARSSGRLVAIKSIRKDKIKOEODILHIREIETMSSLNPHI	112
QY	61	IAIHVEFPENSKTIIVMEYASRGDLYDIYSESRPKSERDARHFFROIYSALFYCHONGIV	120
DB	113	IAIHVEFPENSKTIIVMEYASRGDLYDIYSESGQLSERARHFFROIYSAYHYCHONRV	172
QY	121	HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYGPEVD	180
DB	173	HRDLKLENIILLDANGNIKIADFGLSNLYHGKFLOTFCGSPLYASPEIYNGKPYGPEVD	232
QY	181	SWSLGLVLLIYILVHGMPFGDGHKTLVQISNGAAREPKSPSDAGCLRILMLNVPTRRA	240
DB	233	SWSLGLVLLIYILVHGMPFGDGHKTLVQISNGAAREPKSPSDAGCLRILMLNVPTRRA	292
QY	241	TLEDVYASHMWNVW	253
DB	293	TLEDVYASHMWNVW	305

```

US-10-423-543-44
; Sequence 44, Application US/10423543
; Publication No. US20040058355A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Libermann, Rosana K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Miyoung
; APPLICANT: Williamson, Mark J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Bandaru, Rajasekhara
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21523, 26176, 26343, 56638,
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
; FILE REFERENCE: MP103-0230MINI
; CURRENT APPLICATION NUMBER: US/10/423,543
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 10/278,036
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 09/711,216
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/205,447
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 10/012,055
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/248,325
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 10/003,690
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/248,893
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 10/217,168
; PRIOR FILING DATE: 2002-08-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-423-543-44

```

Query Match	96.3%;	Score 1293;	DB 15;	Length 628;
Best Local Similarity	95.7%;	Pred. No. 1,le-96;		
Matches	242;	Conservative	6;	Mismatches 5;
			Indels	0;
			Gaps	0;
QY	1	YEPLTLAGGYGKVKARSSGRLVAIKSIRKDKIKOEODLLHRIREIETMSLNHPHI	60	
Db	53	YEPLTLAGGYGKVKARSSGRLVAIKSIRKDKIKOEODLLHRIREIETMSLNHPHI	112	
QY	61	IAIHVEFNSSKIVIVMEYASRGDLYDYSERPRISERDARHPFOIVSAALHYCHONGIV	120	
Db	113	IAIHVEFNSSKIVIVMEYASRGDLYDYSERQDLSERARARFPFOIVSAVHICHONRV	172	
QY	121	HRDLKLENILLDANGNIKIADFGLSNLVHKGFLOTFCGSPDYASPEIVNGKPYVPEVD	180	
Db	173	HRDLKLENILLDANGNIKIADFGLSNLVHGKFLQTFCGSPDYASPEIVNGKPYTPEVD	232	
QY	181	SMSLGVLLYIIVHGMPDPGDQDHTLVQVISNGAARREPKSDACGLIHWLLMVNPTTRA	240	
Db	233	SMSLGVLLYIIVHGMPDPGDGHKLVQVISNGAARREPKSDACGLIHWLLMVNPTTRA	292	

```

QY      241 TLEDVASHMWVM 253
      293 TLEDVASHMWVM 305

RESULT 7
US-10-618-941-77
; Sequence 77, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERRARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-77

Query Match      96.3%; Score 1293; DB 16; Length 672;
Best Local Similarity 95.7%; Pred. No. 1.2e-96;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      1 YEFLETLGKGYGKVKKARESSGRIVAISIRKDKIKDEODLHIREIEMSSLNHPHI 60
      97 YEFLETLGKGYGKVKKARESSGRIVAISIRKDKIKDEODLHIREIEMSSLNHPHI 156
QY      61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRFOIVSALHYCHONGIV 120
      157 IAIHEVFENSSKIVIMEYASRGDLVDYISERQOLSERERARHFRFOIVSALHYCHONGIV 216
Db      121 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
      217 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 276
QY      181 SMSLGLVLIYLVHGTMPPFDGQDHKTIVKQISNGAYREPPKPSDAGCLRMLMNVPTTRA 240
      277 SMSLGLVLIYLVHGTMPPFDGQDHKTIVKQISNGAYREPPKPSDAGCLRMLMNVPTTRA 336
Db      241 TLEDVASHMWVM 253
      337 TLEDVASHMWVM 349

RESULT 8
US-10-311-034-12
; Sequence 12, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROMSKI, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Aameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam

```

```

; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAPALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: DING, Li
; APPLICANT: GREYER, Megan
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: BATRA, Sajeiv
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
; CURRENT APPLICATION NUMBER: US/10/311,034
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372,
; 60/228,056
; PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
; 25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CD1
US-10-311-034-12

Query Match      95.8%; Score 1286; DB 15; Length 594;
Best Local Similarity 95.3%; Pred. No. 3.9e-96;
Matches 241; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      1 YEFLETLGKGYGKVKKARESSGRIVAISIRKDKIKDEODLHIREIEMSSLNHPHI 60
      19 YEFLETLGKGYGKVKKARESSGRIVAISIRKDKIKDEODLHIREIEMSSLNHPHI 78
Db      61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRFOIVSALHYCHONGIV 120
      79 IAIHEVFENSSKIVIMEYASRGDLVDYISERQOLSERERARHFRFOIVSALHYCHONGIV 138
QY      121 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
      139 HRDLKLENIILDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 198
Db      181 SMSLGLVLIYLVHGTMPPFDGQDHKTIVKQISNGAYREPPKPSDAGCLRMLMNVPTTRA 240
      199 SMSLGLVLIYLVHGTMPPFDGQDHKTIVKQISNGAYREPPKPSDAGCLRMLMNVPTTRA 258
QY      241 TLEDVASHMWVM 253
      259 TLEDVASHMWVM 271

RESULT 9
US-10-343-514-101
; Sequence 101, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFSBARE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK

```

FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 101
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-10-343-514-101

Query Match 95.1%; Score 1276; DB 16; Length 251;
Best Local Similarity 95.6%; Pred. No. 9.8e-96;
Matches 240; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 60
DB 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 60
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPRLSERDARHPROQVSAHYCHONGIV 120
DB 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERQOLSREKRRHFFRQIVSAHYCHONGIV 120
QY 121 HRDLKLENIILLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYVNGKPYGPEVD 180
DB 121 HRDLKLENIILLDANGNIKIADFGLSNLVHOGKFLQTFCSPLVASPEIYVNGKPYTGPEDV 180
QY 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
QY 241 TLEDVASHMWV 251
DB 241 TLEDVASHMWV 251

RESULT 10

US-10-322-281-23
Sequence 23, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 640
TYPE: PRT
ORGANISM: Mus musculus
US-10-322-281-23

Query Match 94.9%; Score 1273; DB 16; Length 640;
Best Local Similarity 97.2%; Pred. No. 4.8e-95;
Matches 246; Conservative 0; Mismatches 1; Indels 6; Gaps 3;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 60
DB 72 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 131
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPRLSERDARHPROQVSAHYCHONGIV 120

DB 132 IAIHEVEFENSSKIVIMEVYASRGDLVDYIS--PRLS--DARHPROQVSAHYCHONGIV 187
QY 121 HRDLKLENIILLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYVNGKPYGPEVD 180
DB 188 HRDLKLENIILLDANGNIKIADFGLSNLVHOGKFLQTFCSPLVASPEIYVNGKPYGPEVD 245
QY 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 246 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 305
QY 241 TLEDVASHMWV 253
DB 306 TLEDVASHMWV 318

RESULT 11

US-10-322-281-26
Sequence 26, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 616
TYPE: PRT
ORGANISM: Homo sapiens
US-10-322-281-26

Query Match 91.7%; Score 1230; DB 16; Length 616;
Best Local Similarity 93.3%; Pred. No. 1.4e-91;
Matches 236; Conservative 6; Mismatches 5; Indels 6; Gaps 3;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 60
DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIDEDDLHIREIIMSLNPHI 112
QY 61 IAIHEVEFENSSKIVIMEVYASRGDLVDYISERPRLSERDARHPROQVSAHYCHONGIV 120
DB 113 IAIHEVEFENSSKIVIMEVYASRGDLVDYISQ--QLSE--ARHPROQVSAHYCHONGIV 168
QY 121 HRDLKLENIILLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYVNGKPYGPEVD 180
DB 169 HRDLKLENIILLDANGNIKIADFGLSNLVHOGKFLQTFCSPLVASPEIYVNGKPYTGPEDV 226
QY 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 240
DB 227 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYREPPKPSDAGLIRWLMVNPTRRA 286
QY 241 TLEDVASHMWV 253
DB 287 TLEDVASHMWV 299

RESULT 12

US-09-780-949-2
Sequence 2, Application US/09780949
Patent No. US20020006618A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Liberman, Rosana
APPLICANT: Welch, Nadine S.
APPLICANT: Galvin, Katherine M.
TITLE OF INVENTION: Methods for Using 20893, a Human Protein
FILE REFERENCE: 035800/209015
CURRENT APPLICATION NUMBER: US/09/780,949
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,690

;; PRIOR FILING DATE: 2000-02-09
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 661
;; TYPE: PRT
;; ORGANISM: H. sapiens
US-09-780-949-2

Query Match 84.1%; Score 1128.5; DB 9; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGYGKVKKARES-SGRVVAIKSIRKDKIKDEODLHTRREIEMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYGKVKRATERFSGRVVAIKSIRKDKIKDEODVHTRREIEMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSISKIVIMEYASRGDLYDIYSERPRLSERDARHFFQIYSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASKGELYDIYSERRRISERETRHFFQIYSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 VHRDLKLENTLLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGRPYVGPV 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 175 VHRDLKLENTLLDNCNICKIADFGLSNLYQKDFLOTFCGSPLYASPEIYNGRPYGPV 234
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 180 DSWSLGVLIIYLVHGTMPDPGQDKTLVKQISNGATREPKPSDAGLIRMLMNPTRR 239
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 235 DSWALGVLLYTLVYGTMPDPGFPHKMLIRQISSGEYRELPQSDAGLIRMLMNPDRR 294
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 ATLEDVASHMWVMV 253
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 295 ATLEDIANHWVMV 308
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
US-09-780-949-6
;; Sequence 6; Application US/09780949
;; Patent No. US20020006618x1
;; GENERAL INFORMATION:
;; APPLICANT: Kapeller-Liberman, Rosana
;; APPLICANT: Welch, Nadine S.
;; APPLICANT: Galvin, Katherine M.
;; TITLE OF INVENTION: Methods for Using 20893, a Human Protein
;; TITLE OF INVENTION: Kinase
;; FILE REFERENCE: 035800/209015
;; CURRENT APPLICATION NUMBER: US/09/780,949
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: US 60/181,690
;; PRIOR FILING DATE: 2000-02-09
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 661
;; TYPE: PRT
;; ORGANISM: H. sapiens
US-09-780-949-6

Query Match 84.1%; Score 1128.5; DB 9; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGYGKVKKARES-SGRVVAIKSIRKDKIKDEODLHTRREIEMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYGKVKRATERFSGRVVAIKSIRKDKIKDEODVHTRREIEMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSISKIVIMEYASRGDLYDIYSERPRLSERDARHFFQIYSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASKGELYDIYSERRRISERETRHFFQIYSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 VHRDLKLENTLLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGRPYVGPV 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 175 VHRDLKLENTLLDNCNICKIADFGLSNLYQKDFLOTFCGSPLYASPEIYNGRPYGPV 234
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 180 DSWSLGVLIIYLVHGTMPDPGQDKTLVKQISNGATREPKPSDAGLIRMLMNPTRR 239
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 235 DSWALGVLLYTLVYGTMPDPGFPHKMLIRQISSGEYRELPQSDAGLIRMLMNPDRR 294
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 ATLEDVASHMWVMV 253
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 295 ATLEDIANHWVMV 308
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
US-10-354-358-82
;; Sequence 82; Application US/10354358
;; Publication No. US20030157082A1
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc
;; APPLICANT: Hunter, John Joseph
;; APPLICANT: MacBeth, Kyle J.
;; APPLICANT: Teal, Fong-Ying
;; APPLICANT: Lesoon, Andrea
;; APPLICANT: Lightcap, Eric S.
;; APPLICANT: Williamson, Mark
;; APPLICANT: Rudolph-Owen, Laura A.
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
;; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
;; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
;; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
;; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
;; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
;; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
;; FILE REFERENCE: MP102-020PRNOMNTM
;; CURRENT APPLICATION NUMBER: US/10/354,358
;; PRIOR FILING DATE: 2003-01-30
;; PRIOR APPLICATION NUMBER: US 60/353,600
;; PRIOR FILING DATE: 2002-01-31
;; PRIOR APPLICATION NUMBER: US 60/364,517
;; PRIOR FILING DATE: 2002-03-15
;; PRIOR APPLICATION NUMBER: US 60/371,075
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: US 60/371,507
;; PRIOR FILING DATE: 2002-04-10
;; PRIOR APPLICATION NUMBER: US 60/372,984
;; PRIOR FILING DATE: 2002-04-16
;; PRIOR APPLICATION NUMBER: US 60/374,194
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: US 60/382,995
;; PRIOR FILING DATE: 2002-05-24
;; PRIOR APPLICATION NUMBER: US 60/385,023
;; PRIOR FILING DATE: 2002-05-31
;; PRIOR APPLICATION NUMBER: US 60/388,853
;; PRIOR FILING DATE: 2002-06-14
;; PRIOR APPLICATION NUMBER: US 60/389,395
;; PRIOR FILING DATE: 2002-06-17
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 122
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 82
;; LENGTH: 661
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-354-358-82

Query Match 84.1%; Score 1128.5; DB 14; Length 661;
Best Local Similarity 82.3%; Pred. No. 2.8e-83;
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETLGKGYGKVKKARES-SGRVVAIKSIRKDKIKDEODLHTRREIEMSLNHPH 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 55 YELQETLGGTGYGKVKRATERFSGRVVAIKSIRKDKIKDEODVHTRREIEMSLNHPH 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 IIAIHVEFENSISKIVIMEYASRGDLYDIYSERPRLSERDARHFFQIYSAVHYCHONGI 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 115 IISIVFENKDKIVIMEYASKGELYDIYSERRRISERETRHFFQIYSAVHYCHONGV 174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

QY 120 VHRDLKLENIILLDANGNIKIADFGISNLVHKGFLOTFCGSPLYASPEIYNGKPYVGPV 179
    |||||
DB 175 VHRDLKLENIILLDANGNIKIADFGISNLVYKOKFLQTFGSPLYASPEIYNGRPRYGPV 234
    |||||
QY 180 DSWMSIGVLLIYLVHGTMPDPGQDHKTLYVQISNGAYREPPKPSDACGLIRMLMNPTR 239
    |||||
DB 235 DSWALGVLLIYLVYGTMPDPGFDHKNLIRQISSGEYREPTQPSDARGLIRMLMNPDRR 294
    |||||
QY 240 ATLEDVASHMWV 253
    |||||
DB 295 ATLEDIANHNMV 308
    |||||

```

```

RESULT 15
US-10-343-514-102
; Sequence 102, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/ECT
; CURRENT APPLICATION NUMBER: US/10/343,514
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 102
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-514-102

```

```

Query Match 82.8%; Score 1111.5; DB 16; Length 252;
Best Local Similarity 82.1%; Pred. No. 2.3e-82;
Matches 207; Conservative 24; Mismatches 20; Indels 1; Gaps 1;

QY 1 YEFLETGKGTGKVKKARS-SGRVLAISKIRKDKIKDBODLHTRREIEMSLNHPH 59
    |||||
DB 1 YELQETLGGTGYKVRATERPSGRVVAISKIKDKIKDBODVHTRREIEMSLNHPH 60
    |||||
QY 60 IIAIHEVFENSSKIVIVMEYASRGDLVDYISERPRLSERDARHFPQIVSAHYCHONGI 119
    |||||
DB 61 IISIEVFENKDKIVIMBYASKGLYDIISERRLSERETRHFPQIVSAHYCHONGV 120
    |||||
QY 120 VHRDLKLENIILLDANGNIKIADFGISNLVHKGFLOTFCGSPLYASPEIYNGKPYVGPV 179
    |||||
DB 121 VHRDLKLENIILLDANGNIKIADFGISNLVYKOKFLQTFGSPLYASPEIYNGRPRYGPV 180
    |||||
QY 180 DSWMSIGVLLIYLVHGTMPDPGQDHKTLYVQISNGAYREPPKPSDACGLIRMLMNPTR 239
    |||||
DB 181 DSWALGVLLIYLVYGTMPDPGFDHKNLIRQISSGEYREPTQPSDARGLIRMLMNPDRR 240
    |||||
QY 240 ATLEDVASHMWV 251
    |||||
DB 241 ATLEDIANHNMV 252
    |||||

```

Search completed: May 11, 2005, 14:42:21
Job time : 53.0882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:09:18 ; Search time 66.3982 Seconds
(without alignments)
1473.690 Million cell updates/sec

```

Title:      US-09-980-464-11_COPY_57_309
Perfect score: 1342
Sequence:   1 YEPLETLGKGTGVKKARE.....VNPTRRATLEDVASHWVNM 253

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

```
Database :
1_ * Genesegp_16Dec04.*
1_ * Genesegp_1980s.*
2_ * Genesegp_1990s.*
3_ * Genesegp_2000s.*
4_ * Genesegp_2001s.*
5_ * Genesegp_2002s.*
6_ * Genesegp_2003s.*
7_ * Genesegp_2003bs.*
8_ * Genesegp_2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1342	100.0	631	1	AAB50056	Aab50056 Murine Ig
2	1342	100.0	652	8	ABO84756	AbO84756 Murine c6
3	1313	97.8	630	5	AAE19885	Aae19885 Rat SNF1/
4	1293	96.3	628	4	AAm93360	Aam93360 Human pol
5	1293	96.3	628	4	ABU53319	Abu53319 Human cel
6	1293	96.3	628	5	ABP69116	Abp69116 Human pol
7	1293	96.3	628	5	AAU79652	Aau79652 Human pro
8	1293	96.3	628	7	ADF76965	Adf76965 Novel hun
9	1293	96.3	628	8	ADL30886	Adl30886 Human pro
10	1293	96.3	628	8	ADL25362	Adl25362 Human SNA
11	1293	96.3	628	8	ADL41461	Adl41461 Novel hun
12	1293	96.3	628	8	ADQ20172	Ado20172 Human PR
13	1293	96.3	628	8	ABO84757	AbO84757 Human car
14	1293	96.3	672	8	ADJ96620	Adj96620 Human cat
15	1286	95.8	594	5	AAE16266	Aae16266 Human Klt
16	1279.5	95.3	611	6	ABP96085	Abp96085 Human pro
17	1268.5	94.5	629	4	AAB65632	Aab65632 Novel
18	1268.5	94.5	629	8	ADI29239	Adi29239 Human MAR
19	1128.5	84.1	660	7	ADN95766	Adn95766 Human BGC
20	1128.5	84.1	661	4	AAE07847	Aae07847 Human p
21	1128.5	84.1	661	4	AAE07846	Aae07846 Human p
22	1128.5	84.1	661	7	ADBE3421	Adbe3421 Human pr
23	1128.5	84.1	661	8	ADJ75331	Adj75331 Marker g
24	1128.5	84.1	661	8	ADJ25353	Adj25353 Human AKR
25	1128.5	84.1	661	8	ADQ19734	Adq19734 Human sol

ALIGNMENTS

45	660.5	49.2	688	7	ADG91727
44	660.5	49.2	688	6	AAE33555
43	660.5	49.2	688	7	ADG91726
42	660.5	49.2	688	6	AAE33554
41	660.5	49.2	660	5	ABBO0432
40	661.5	49.3	752	8	ADJ96622
39	665.5	49.6	752	8	ABBE2866
38	665.5	49.6	752	8	ADL31225
37	665.5	49.6	752	5	ABG73194
36	665.5	49.6	752	5	ABE04433
35	665.5	49.6	752	4	AAE11782
34	665.5	49.6	752	4	AAE13956
33	665.5	49.6	689	7	ADPF4129
32	665.5	49.6	688	5	ABBO0434
31	665.5	49.6	688	5	ADL29233
30	665.5	49.6	520	8	AD129235
29	665.5	49.6	508	4	AAE65629
28	814	60.7	530	4	AAE00668
27	822.5	61.3	434	4	AAE67451
26	1085	80.8	534	4	AAE71959
25					Abb71959
24					Abbf64755
23					Abb67495
22					Abb67495
21					Abb67495
20					Abb67495
19					Abb67495
18					Abb67495
17					Abb67495
16					Abb67495
15					Abb67495
14					Abb67495
13					Abb67495
12					Abb67495
11					Abb67495
10					Abb67495
9					Abb67495
8					Abb67495
7					Abb67495
6					Abb67495
5					Abb67495
4					Abb67495
3					Abb67495
2					Abb67495
1					Abb67495

RESULT 1	
AAB50056	
ID	AAB50056 standard; protein; 631 AA

DT 19-MAR-2001 (first entry)

DE Murine Lymph node Stromal cell kinase 1.

KW Murine, lymph node Stromal cell kinase; MLK-1; autoimmune disorder;
 KW wound healing; periodontal disease; inflammatory disease; tumour;
 KW infection; allergy.

OS Mus musculus.

PN WO200073468-A1.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US014696.

PR 28-MAY-1999; 99US-0136781P.

PA (IMMV) IMMUNEX CORP.

PI Bird TA, Virca GD, Martin U, Anderson DM;

DR WPI; 2001-061546/07.

2. \times \times

PT Novel murine and human kinase nucleic acids useful for treating inflammations, infections, tumors, allergies, autoimmune diseases, and for stimulating or suppressing immune responses.

PS Claim 10; Page 94-96; 106pp; English.

CC The present sequence is Murine Lymph node Stromal cell kinase 1 (MLSK-1).
CC This protein is useful for treating a variety of disorders listed in the
CC disclosure of the specification, including autoimmune disorders, allergic
CC reactions, myeloid or lymphoid cell deficiencies, wound healing, and
CC tissue repair and replacement, burns, incisions and ulcers, periodontal
CC disease, inflammatory diseases, tumours and bacterial, viral or fungal
CC infection

SQ	Sequence	631	AA;
Query Match	100.0%;	Score 1342;	DB 4; Length 631;

Best Local Similarity 100.0%; Pred. No. 2.7e-138;

Qy	1	YEFLETLKGTYGKXKKAESSGRVVAISIRDKTKDEBDLHIREIIMSLSNHPH	60
Db	57	YEFLETLGTYGKXKKAESSGRVVAISIRDKTKDEBDLHIREIIMSLSNHPH	116
Qy	61	IAIHVEFENSXIIVIMEYASRGDLVDYISERPLSERDRAHFPQVSLAHYCHONGIV	120
Db	117	IAIHVEFENSXIIVIMEYASRGDLVDYISERPLSERDRAHFPQVSLAHYCHONGIV	176
Qy	121	HRDLKLENIILLDANGNIKIADPGLSNLHYHKGKTLQTPCGSPLVYASPEIYNGKPYGVEVD	180
Db	177	HRDLKLENIILLDANGNIKIADPGLSNLHYHKGKTLQTPCGSPLVYASPEIYNGKPYGVEVD	236
Qy	181	SMSLGLVLIYLIYHGTMPFGODHKTLVKOISNAYAREPMPSPDAGCIRWLLMVNPPRRA	240
Db	237	SMSLGLVLIYLIYHGTMPFGODHKTLVKOISNAYAREPMPSPDAGCIRWLLMVNPPRRA	296
Qy	241	TLEDVASHMWVNW	253
Db	297	TLEDVASHMWVNW	309

RESULT 2

ID	AB084756 standard; protein; 652 AA.
XX	
XX	AB084756;
AC	
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Murine cancer-associated protein (CAP) MP07-004.
XX	
KW	Mouse; cancer-associated protein; CAP; cancer; cytostatic.
XX	
XX	Mus musculus.
OS	
XX	
PN	WO2004058146-A2.
XX	
PD	15-JUL-2004.
XX	
PF	15-DEC-2003; 2003WO-US040081.
XX	
PR	17-DEC-2002; 2002US-00322281.
XX	
PA	(SAGR-) SAGRES DISCOVERY INC.
XX	
PI	Morris DW, Malandro MS;
XX	
DR	WPI; 2004-499109/47.
DR	N-PSDB; ABD33083.
XX	
PT	Novel human cancer associated protein encoded within open reading frame
PT	of cancer associated gene, useful as targets for diagnosing cancer.
XX	
PS	Disclosure; SEQ ID NO 23; 182bp; English.
XX	
CC	The invention relates to cancer-associated proteins (CAP) and the cancer-
CC	associated (CA) nucleic acids encoding them. The invention also relates
CC	to a method for treating cancers involving administering to a patient an
CC	inhibitor of CAP, and a method of screening for anticancer activity in a
CC	potential drug involving providing a cell that expresses a CA gene,
CC	contacting a tissue sample derived from a cancer cell with an anticancer
CC	drug candidate and monitoring the effect of the anticancer drug candidate
CC	on expression of the CA gene. The CAP proteins are useful for detecting
CC	cancer associated with expression of a CAP protein in a test cell sample
CC	and for screening for a bioactive agent capable of modulating the
CC	activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC	cancer, involving determining the expression of a CA nucleic acid in a
CC	tissue. This sequence represents a murine CAP of the invention. Note: The
CC	sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO

CC at fcp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 652 AA;

Query Match	100.0%;	Score 1342;	DB 8;	Length 652;
Best Local Similarity	100.0%;	Pred. No. 2.8e-18;		
Matches 253; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	IEPFLTLTAKGYKAKKARSSGRVLAIKSIRKOKIKBOBODLHTRRELEIMSSLNHPH	60
Db	78	YEPFLTLKNGYKAKKARSSGRVLAIKSIRKOKIKBOBODLHTRRELEIMSSLNHPH	137
QY	61	IAIHVEFENSCKIIVMEVYASRGDLUDYISERPRISERDARHFFROIVSAIYHCHONGIV	120
Db	138	IAIHVEFENSCKIIVMEVYASRGDLUDYISERPRISERDARHFFROIVSAIYHCHONGIV	197
QY	121	HRDLKLEINILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASBEIYNGKPYGPEVD	180
Db	198	HRDLKLEINILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASBEIYNGKPYGPEVD	257
QY	181	SMSLGLVLLYIIVHGMFPDGDGDKTLVYKQISNGAIREPPKPSDACGLIRMLMVNPTBRA	240
Db	258	SMSLGLVLLYIIVHGMFPDGDGDKTLVYKQISNGAIREPPKPSDACGLIRMLMVNPTBRA	317
QY	241	TLEDVASHMWNWV 253	
Db	318	TLEDVASHMWNWV 330	

RESULT 3

ID	AAE19885	standard; protein; 630 AA.
XX		
AC	AAE19885;	
XX		
DT	18-JUN-2002	(first entry)
XX		
DE	Rat SNF1/AMPK-Related Kinase (SNARK) protein.	
XX		
KM	Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;	
KM	glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;	
KM	hyperglycaemic; drug screening; hypoglycaemia.	
XX		
OS	Rattus sp.	
XX		
PH	Key	Location/Qualifiers
FT	Binding-site	63..89
FT		/note="Protein kinase ATP-binding region signature"
FT	Domain	137..140
FT		/note="Serine/threonine kinase catalytic domain"
FT	Active-site	175..187
FT		/note="Serine/threonine protein kinase active-site"
FT	Domain	297..300
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	335..338
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	381..384
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	422..425
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	468..471
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	517..520
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	601..604
FT		/note="Serine/threonine kinase catalytic domain"
FT	Domain	608..611
FT		/note="Serine/threonine kinase catalytic domain"
XX		
FN	W0200212456-A2.	
XX		
DD	14-FEB-2002.	
XX		

PF 02-AUG-2001; 2001MO-CA001109.
XX
XX 03-AUG-2000; 2000US-0222650P.
PR 12-MAR-2001; 2001US-0274613P.
PR 26-MAR-2001; 2001CA-02340783.
XX
XX (ONBO-) 1149336 ONTARIO INC.
XX
XX Drucker DJ, Rosen CF, Lefebvre DL,
PI
XX
XX WPI: 2002-241747/29.
DR N-PSDB; AAD31710.
XX
XX AMPK (AMP-activated protein kinase)-related kinase, designated SNARK
PT polypeptides and polynucleotides, useful for treating or preventing
PT diabetes, or other disorders of lipoprotein production leading to
XX increased levels of cholesterol.
XX
XX Claim 1; Fig 2; 94pp; English.
XX
XX The invention relates to an AMPK (AMP-activated protein kinase)-related
CC kinase, designated SNARK polypeptides and polynucleotides. SNARK
CC (SNIT/AMP-activated protein kinase) is involved in stress response to
CC glucose deprivation. The polynucleotides are useful for expressing SNARK
CC protein in isolated form or as a protein conjugate. Activation of SNARK
CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
CC in other cell types such as heart and skeletal muscles, as well as
CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
CC to have insulin-like effects that would enhance the disposal of glucose
CC into muscle and reduce plasma glucose for the treatment of diabetes and
CC some type of disorder of lipoprotein production leading to increased
CC levels of cholesterol or triglycerides. SNARK or its variants may be
CC administered to a subject to treat or prevent a disease associated with
CC decreased expression of SNARK, such as diabetes. SNARK antibodies are
CC used to modulate SNARK activity either in vivo for therapeutic purposes,
CC or in vitro, for drug screening and related investigational purposes.
CC SNARK antagonists may be administered to increase fuel production,
CC decrease glucose uptake and increase levels of blood glucose in a patient
CC suffering from hypoglycaemia. The present sequence is rat SNARK protein
XX
XX Sequence 630 AA;
SQ

Query Match 97.8%; Score 1313; DB 5; Length 630;
Best Local Similarity 98.0%; Pred. No. 4.3e-135;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 116
QY 61 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
DB 117 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 176
QY 121 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 180
DB 177 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 236
QY 181 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYRPPKPSDAGLIRWLAVNPTRR 240
DB 237 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYRPPKPSDAGLIRWLAVNPTRR 296
QY 241 TLEDVASHMWVMV 253
DB 297 TLEDVASHMWVMV 309

RESULT 4
ID AAM93360 standard; protein; 628 AA.
XX
XX AAM93360;
XX

DT 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 2919.
DE
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
OS
XX EP1130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-00114089.
PF
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Nishikawa T, Iecgai T, Hayashi K, Ishii S, Kawai Y,
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94280.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 2919; 1380pp + Sequence Listing; English.
PS
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
XX Sequence 628 AA;
SQ

Query Match 96.3%; Score 1293; DB 4; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 60
DB 53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNHPH 112
QY 61 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120
DB 113 IAHHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 172
QY 121 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 180
DB 173 HRDLKLENTILLDANGNIKIADFGLSNLVHKGFLOTFCCSPLYASPEIYVNGKPYGPEVD 232
QY 181 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYRPPKPSDAGLIRWLAVNPTRR 240
DB 233 SMSLGLVLLYLVHGTMPPDQGDHKTIVKQISNGAYRPPKPSDAGLIRWLAVNPTRR 292
QY 241 TLEDVASHMWVMV 253
DB 293 TLEDVASHMWVMV 305

RESULT 5
ABUS3319

```
ID      ABU5319 standard; protein; 628 AA.
XX
AC      ABU5319;
XX
DT      14-APR-2003 (first entry)
XX
DE      Human cell cycle-associated protein from DKFzphes3_7j3.
XX
KM      Human; gene therapy; vaccine; disease treatment; detection.
XX
OS      Homo sapiens.
XX
PN      WO200112659-A2.
XX
PD      22-FEB-2001.
XX
PF      18-AUG-2000; 2000WO-IB001496.
XX
PR      18-AUG-1999; 99US-0149499P.
XX
PR      28-SEP-1999; 99US-0156503P.
XX
PA      (GENU-) GERMAN HUMAN GENOME PROJECT.
XX
PI      Wiemann S;
XX
PI      WPI; 2001-327840/34.
XX
DR      N-PSDB; ABX71420.
XX
PT      Nucleic acids having the sequences of clones isolated from libraries of
XX      different human tissues, useful in recombinant DNA methodologies.
XX
PS      Claim 21; Page 943; 1095pp; English.
XX
CC      This invention describes novel polynucleotides and polypeptides isolated
XX      from human cDNA libraries which can be used for gene therapy or in
XX      vaccines. The polynucleotides of the invention and antibodies encoded by
XX      them may be used in the prevention, diagnosis and treatment of diseases
XX      associated with inappropriate polypeptide expression. The products of the
XX      invention may also be used to identify modulators of expression and
XX      activity and to down regulate expression and activity. The antibodies of
XX      the invention may also be used as diagnostic agents for detecting the
XX      presence of polypeptides in samples. This sequence represents a
XX      polypeptide described in the disclosure of the invention
XX
SQ      Sequence 628 AA;
XX
Query Match      96.3%; Score 1293; DB 4; Length 628;
Best Local Similarity 95.7%; Pred. No. 6,8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
OY      1 YEFLETLGKGTGKTKKARESSGRVLAISKIKDKIKEDODLHIREIFIMSLNHPHI 60
DB      53 YEFLETLGKGTGKTKKARESSGRVLAISKIKDKIKEDODLHIREIFIMSLNHPHI 112
OY      61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHPFOIVSAVHYCHONGIV 120
DB      113 IAIHEVFENSSKIIVMEYASRGDLVDYISERQULSEREAKHFFRQIVSAVHYCHONGIV 172
OY      121 HEDLKIENILLDANGNIKIADFGLSNLYHKGKFLQTFCCGSLYASPEIYVNGKPYGPEVD 180
DB      173 HEDLKIENILLDANGNIKIADFGLSNLYHKGKFLQTFCCGSLYASPEIYVNGKPYGPEVD 232
OY      181 SMSLGLVLIYLVHGTMPPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLVMVNPTRRA 240
DB      233 SMSLGLVLIYLVHGTMPPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLVMVNPTRRA 292
OY      241 TLEDVASHMMVMNM 253
DB      293 TLEDVASHMMVMNM 305
RESULT 6
ABP69116
```

```
ID      ABP69116 standard; protein; 628 AA.
XX
AC      ABP69116;
XX
DT      20-JAN-2003 (first entry)
XX
DE      Human polypeptide SEQ ID NO 1163.
XX
KM      Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX      cell-proliferative disorder; neurodegenerative disease; bacterial;
XX      Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX      multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX      arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
XX      antiparkinsonian; antididiabetic; immunosuppressive; dermatological;
XX      haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
XX      antithratic.
XX
OS      Homo sapiens.
XX
PN      WO200270539-A2.
XX
PD      12-SEP-2002.
XX
PF      05-MAR-2002; 2002WO-US005095.
XX
PR      05-MAR-2001; 2001US-00799451.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Zhou P, Goodrich RW, Aundt V, Zhang J, Zhao QA, Ren F;
XX      Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
XX      Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR      WPI; 2002-759812/82.
XX
DR      N-PSDB; ABZ11333.
XX
PT      New polynucleotides comprising sequences assembled from expressed
XX      sequence tags (ESTs), useful for treating cell-proliferative,
XX      neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
XX      or coagulation disorders.
XX
PS      Claim 9; SEQ ID NO 1163; 1012pp + Sequence Listing; English.
XX
CC      The invention relates to an isolated polynucleotide (1) comprising a
XX      nucleotide sequence selected from any of 948 sequences (ABZ1119-
XX      ABZ12066) or their mature protein coding portion, active domain coding
XX      protein or complementary sequences. The polynucleotides are useful for
XX      identifying expressed genes or for physical mapping of human genome. The
XX      encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
XX      markers, as a food supplement, for generating antibodies, in medical
XX      imaging, screening and diagnostic assays and for treating cell-
XX      proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
XX      or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
XX      diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
XX      platelet or coagulation disorders, wound, burns, incision, ulcers, liver
XX      or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
XX      arthritis, etc. Note: The sequence data for this patent did not form part
XX      of the printed specification, but was obtained in electronic format
XX      directly from WIPO at ftp.wipo.int/pub/published_pot_sequences
XX
SQ      Sequence 628 AA;
XX
Query Match      96.3%; Score 1293; DB 5; Length 628;
Best Local Similarity 95.7%; Pred. No. 6,8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
OY      1 YEFLETLGKGTGKTKKARESSGRVLAISKIKDKIKEDODLHIREIFIMSLNHPHI 60
DB      53 YEFLETLGKGTGKTKKARESSGRVLAISKIKDKIKEDODLHIREIFIMSLNHPHI 112
OY      61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHPFOIVSAVHYCHONGIV 120
DB      113 IAIHEVFENSSKIIVMEYASRGDLVDYISERQULSEREAKHFFRQIVSAVHYCHONGIV 172
```

QY 121 HRDLKLENTLLDANGNIKIADFGSLNLYHKFKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
 CC |||||
 XX |||||
 Db 173 HRDLKLENTLLDANGNIKIADFGSLNLYHQGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
 CC |||||
 QY 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYRPPKPSDAGLIRMLMVPNPTTRA 240
 CC |||||
 Db 233 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYRPPKPSDAGLIRMLMVPNPTTRA 292
 CC |||||
 QY 241 TLEDVASHMWVMW 253
 CC |||||
 Db 293 TLEDVASHMWVMW 305
 CC |||||
 RESULT 7
 AAU79652
 ID AAU79652 standard; protein; 628 AA.
 XX
 AC AAU79652;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human protein kinase 3700.
 XX
 KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
 KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
 KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
 KW cell proliferation disorder; cell differentiation disorder; carcinoma;
 KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;
 KW cytosolic; antiatherosclerotic; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN WO200224921-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US030115.
 XX
 PR 25-SEP-2000; 2000US-0234922P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ, Galvin KM;
 XX
 DR WPI; 2002-352007/38.
 DR N-PSDB; ABK14000.
 XX
 PT Use of modulators of activity of 3700 protein for making medicament for
 PT e.g., modulating protein phosphorylation or cell signaling, or for
 PT treating or preventing cellular proliferative and/or differentiative
 PT disorders.
 XX
 PS Claim 19; Fig 1; 115pp; English.
 XX
 CC The present invention relates to the isolation of a novel human protein
 CC kinase designated 3700, and the polynucleotide sequence encoding it. The
 CC invention also describes the use of a modulator of the activity of
 CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
 CC composition for modulating the ability of a cell to phosphorylate an
 CC amino acid residue of a substrate protein. Modulators of protein kinase
 CC 3700 activity are useful for modulating protein phosphorylation, cell
 CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
 CC angiogenesis, tissue repair, tissue regeneration, establishment or
 CC progression of atherosclerosis, and signalling across the blood-brain
 CC barrier. The polynucleotide and polypeptide molecules for protein kinase
 CC 3700 may be used as diagnostic targets and therapeutic agents for
 CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
 CC curing PK-related disorders and cellular proliferative and/or
 CC differentiative disorders (e.g. haematopoietic neoplastic disorders,
 CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
 CC polynucleotide sequence can be used to express protein kinase 3700, to
 CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for

CC tissue typing, in forensic biology, and as surrogate markers. The present
 CC sequence represents human protein kinase 3700
 XX
 SQ Sequence 628 AA;
 Query Match 96.3%; Score 1293; DB 5; Length 628;
 Best Local Similarity 95.7%; Pred. No. 6, 8e-133;
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 1 YEFLETLGKGTGKVKARESSGRVAIAKIRKDKIKDDDLIRRIEIMSLNPHI 60
 CC |||||
 Db 53 YEFLETLGKGTGKVKARESSGRVAIAKIRKDKIKDDDLIRRIEIMSLNPHI 112
 CC |||||
 QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSEDRARHFRQIVSALHYCHONGIV 120
 CC |||||
 Db 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSERBARHFRQIVSAVHCQGNRVV 172
 CC |||||
 QY 121 HRDLKLENTLLDANGNIKIADFGSLNLYHKFKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
 CC |||||
 Db 173 HRDLKLENTLLDANGNIKIADFGSLNLYHQGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
 CC |||||
 QY 181 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYRPPKPSDAGLIRMLMVPNPTTRA 240
 CC |||||
 Db 233 SMSLGVLLYLVHGTMPPDGDHKTIVKQISNGAYRPPKPSDAGLIRMLMVPNPTTRA 292
 CC |||||
 QY 241 TLEDVASHMWVMW 253
 CC |||||
 Db 293 TLEDVASHMWVMW 305
 CC |||||
 RESULT 8
 ADF76965
 ID ADF76965 standard; protein; 628 AA.
 XX
 AC ADF76965;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Novel human secreted and transmembrane protein Seqid 640.
 XX
 KW human; PRO; membrane bound protein; membrane bound receptor;
 KW cell proliferation; cell migration; cell differentiation;
 KW mitogenic factor; survival factor; cytotoxic factor;
 KW differentiation factor; neuropeptide; hormone; cell receptor;
 KW receptor-ligand interaction; cytosolic; chondrocyte; tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO2003072035-A2.
 XX
 PD 04-SEP-2003.
 XX
 PF 21-FEB-2003; 2003WO-US005241.
 XX
 PR 22-FEB-2002; 2002US-0359461P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR,
 PI Williams PM, Wood WT, Wu TD;
 XX
 DR WPI; 2003-721702/68.
 DR N-PSDB; ADF76964.
 XX
 PT New PRO polypeptides, useful for diagnosing and treating an immune
 PT related disorder, e.g., systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
 PT diabetes mellitus.
 XX
 PS Claim 10; SEQ ID NO 640; 918pp; English.
 XX
 CC This invention relates to novel nucleic acids encoding human PRO secreted
 CC and transmembrane proteins. Extracellular proteins play important roles

CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC stimulation of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is the amino acid sequence of a human PRO protein of the
CC invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 7; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YFFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 60
DB 53 YFFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 112
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHPFRQIVSALHYCHONGIV 120
DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQOLSERARHPFRQIVSAVHYCHONRV 172
QY 121 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
DB 173 HRDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLMVPNTRRA 240
DB 233 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLMVPNTRRA 292
QY 241 TLEDVASHMWVMV 253
DB 293 TLEDVASHMWVMV 305

RESULT 9

ADL30886
ID ADL30886 standard; protein; 628 AA.

AC ADL30886;

DT 20-MAY-2004 (first entry)

DE Human protein encoded by a full length cDNA clone Segid 2919.

KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method.

OS Homo sapiens.

XX EPI396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.
XX 07-JUL-2000; 2000EP-00114089.
XX

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Oca T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
DR N-PSDB; ADL30885.

XX New oligonucleotide primers (830 CDNAs) useful for synthesizing full
PT length human cDNAs.

PS Example 1; SEQ ID NO 2919; 1340bp; English.

CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YFFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 60
DB 53 YFFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPHI 112
QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHPFRQIVSALHYCHONGIV 120
DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQOLSERARHPFRQIVSAVHYCHONRV 172
QY 121 HEDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 180
DB 173 HRDLKIENILLDANGNIKIADFGLSNLVHGKFLQTFCCSPLYASPEIYNGKPYGPEVD 232
QY 181 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLMVPNTRRA 240
DB 233 SMSLGLVLYILVHGMPFGQDHKTLVKQISNGAYRPPKPSDAGLIRMLMVPNTRRA 292
QY 241 TLEDVASHMWVMV 253
DB 293 TLEDVASHMWVMV 305

RESULT 10

ADL25362
ID ADL25362 standard; protein; 628 AA.

AC ADL25362;

DT 03-JUN-2004 (first entry)

DE Human SNARK, SEQ ID 12.

KW Cytostatic; Neuroprotective; Nootropic; Muscular; Gene therapy;
KW Adenosine Monophosphate activated Protein Kinase-related kinase 5;
KW AMPK-related kinase 5; ARK5; KIAA0537; stress resistance; tumour;
KW nervous disorder; muscle disorder; ataxia telangiectasia; SNARK.

OS Homo sapiens.

XX WO2004019994-A1.

XX 11-MAR-2004.

XX 19-AUG-2003; 2003WO-0P010435.
XX

PR 27-AUG-2002; 2002JP-00247761.
XX (NINA-) JAPAN AGENCY NAT INST HEALTH;
PA (PHAR-) ORG PHARM SAFETY & RES.
XX
PI Esumi H, Suzuki A;
XX WPI; 2004-248195/23.
DR
XX Agent for imparting stress resistance to cells, comprises DNA encoding
PT for adenosine monophosphate activated protein kinase-related kinase 5 for
PT treatment of tumors, muscle and nervous disorders, and ataxia
PT teleangiectasia.
PS
XX Example 1; SEQ ID NO 12; 143bp; Japanese.
XX
XX The present invention relates to human Adenosine Monophosphate activated
CC Protein Kinase (AMPK)-related kinase 5 (ARK5/ KIAA0537) and its coding
CC sequence, which are useful for imparting stress resistance to cells. ARK5
CC is useful for treatment and prevention of tumours, nervous disorders,
CC muscle disorders and ataxia teleangiectasia. The present sequence was
CC used to illustrate the invention.
XX
XX
SQ Sequence 628 AA;
Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 YEPLELTGKGTGKVKKARESSGRVLAIKSIRKDKIKEDDLIIRREIEMSLNHPH 60
DB 53 YEFLELTGKGTGKVKKARESSGRVLAIKSIRKDKIKEDDLIIRREIEMSLNHPH 112
QY 61 IAIHEFEKSSKIVIMEYASRGDLYDYSERPRLSERPARHFRQIVGALHYCHONGIV 120
DB 113 IAIHEFEKSSKIVIMEYASRGDLYDYSERQSLSERPRHFRQIVSAVHCHONRVV 172
QY 121 HRDLKLENIILDPANGIKIADFGLSNLYHKGKPLQTFCCSPLYASPEIYNGKPYPEVD 180
DB 173 HRDLKLENIILDPANGIKIADFGLSNLYHQGKPLQTFCCSPLYASPEIYNGKPYPEVD 232
QY 181 SMSLGLVLYLVHGMTPPDGQDHKTLYVKQISNGAYRPPKPSDAGLIRMLMVPTRRA 240
DB 233 SMSLGLVLYLVHGMTPPDGQDHKTLYVKQISNGAYRPPKPSDAGLIRMLMVPTRRA 292
QY 241 TLEDVASHWVNM 253
DB 293 TLEDVASHWVNM 305
RESULT 11
ADL14161
ID ADL14161 standard; protein, 628 AA.
XX
AC ADL14161;
XX
XX 17-JUN-2004 (first entry)
DT
XX Novel human gene 3700 encoded protein.
DE
XX cytostatic; cardiac; hypotensive; antianginal; osteopathic;
KW antiarthritic; antiinflammatory; neuroprotective; antidiabetic;
XX antiapoptotic; antithrombotic; cardiovascular; vitamin; analgesic; CNS;
KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;
KW nephroprotective; antithyroid; dermatological; immunomodulator;
KW cell proliferation disorder; cell differentiation disorder;
KW brain disorder; platelet disorder; breast disorder; colon disorder;
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;
KW thyroid disorder; testes disorder; haematopoietic disorder;
KW pancreatic disorder; skeletal muscle disorder; skin disorder;
KW dermal disorder; bone metabolism disorder; immune disorder;
KW inflammatory disorder; cardiovascular disorder;

KW endothelial cell disorder; liver disorder; viral disease; pain disorder;
KW metabolic disorder; neurological disorder;
KW central nervous system disorder; erythroid disorder;
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
KW cell proliferation; cell differentiation; cell growth; cell division;
KW human.
XX
XX Homo sapiens.
XX
XX US2004058355-A1.
XX
PD 25-MAR-2004.
XX
XX 25-APR-2003; 2003US-00423543.
XX
PR 30-SEP-1998; 98US-00163821.
PR 27-JAN-1993; 99US-0117580P.
PR 25-MAR-1999; 99US-00276400.
PR 30-JUL-1999; 99US-00365162.
PR 09-SEP-1999; 99US-00392189.
PR 05-OCT-1999; 99US-00412210.
PR 23-NOV-1999; 99US-00448076.
PR 29-FEB-2000; 2000US-0186061P.
PR 28-APR-2000; 2000US-0200688P.
PR 19-MAY-2000; 2000US-0205447P.
PR 30-JUN-2000; 2000US-00608921.
PR 31-JUL-2000; 2000US-0221925P.
PR 25-SEP-2000; 2000US-0234322P.
PR 25-SEP-2000; 2000US-0235035P.
PR 08-NOV-2000; 2000US-0246669P.
PR 09-NOV-2000; 2000US-00711216.
PR 14-NOV-2000; 2000US-0248325P.
PR 15-NOV-2000; 2000US-0248893P.
PR 22-DEC-2000; 2000US-0257511P.
PR 05-JAN-2001; 2001US-0260166P.
PR 28-FEB-2001; 2001US-0079703P.
PR 27-APR-2001; 2001US-00845044.
PR 20-JUL-2001; 2001US-00909743.
PR 31-UTL-2001; 2001US-00920346.
PR 13-AUG-2001; 2001US-00928531.
PR 14-AUG-2001; 2001US-00929218.
PR 15-AUG-2001; 2001US-0312539P.
PR 25-SEP-2001; 2001US-00963159.
PR 08-NOV-2001; 2001US-00908016.
PR 13-NOV-2001; 2001US-00012055.
PR 15-NOV-2001; 2001US-00003690.
PR 30-JAN-2002; 2002US-0060763.
PR 25-MAR-2002; 2002US-00105989.
PR 12-APR-2002; 2002US-00121911.
PR 12-AUG-2002; 2002US-00217168.
PR 22-OCT-2002; 2002US-00278036.
PR 02-JAN-2003; 2003US-00336489.
PR 03-JAN-2003; 2003US-00336153.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Kapeller-Libermann R, Hunter JU, Meyers RE, Rudolph-Owen LA;
PI Curtis RAJ, Olandt PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;
PI Siles-Santiago I, Bandaru R;
XX
DR WPI; 2004-268788/25.
DR N-PSDB; ADL14160, ADL14162.
XX
XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,
PT 26176, 26333, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593
PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,
PT heart failure and angina.
XX
PS Claim 4; SEQ ID NO 44; 139pp; English.
XX
XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,

CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (1) comprising
CC any one of 40 nucleotide sequences (1). The nucleic acid molecules and
CC polypeptides are useful for diagnosing and treating a subject having a
CC disorder, or a subject at risk of developing a disorder, which is
CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,
CC m1983, 38555 or 593 activity, such as cellular proliferative and/or
CC differentiative disorders, brain disorders, platelet disorders, breast
CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
CC ovarian disorders, prostate disorders, cervical disorders, spleen
CC disorders, thymus disorders, thyroid disorders, testes disorders,
CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
CC skin (dermal) disorders, disorders associated with bone metabolism,
CC immune, e.g., inflammatory disorders, cardiovascular disorders,
CC endothelial cell disorders, liver disorders, viral diseases, pain
CC disorders, metabolic disorders, neurological or central nervous system
CC disorders, erythroid disorders, blood vessel disorders or angiogenic
CC disorders (all claimed), e.g., cancer, heart failure, hypertension,
CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
CC disease, psoriasis, or asthma. The nucleic acid molecules and
CC polypeptides are also useful as modulating agents in regulating a variety
CC of cellular processes, e.g., cell proliferation, differentiation, growth and
CC division. This is the amino acid sequence of a novel human protein of the
CC invention. Note: The sequences given in the specification are also
CC available in electronic format from
CC ftp.segdata.uspto.gov/Sequence.html?DocID=20040058355.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;

Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGYGKVKARSSGRVLAIKSIRKDKI KDEDLHRRREIIMSSLNHPH 60

DB 53 YEFLETLGKGTGYGKVKARSSGRVLAIKSIRKDKI KDEDLHRRREIIMSSLNHPH 112

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120

DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172

QY 121 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 180

DB 173 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 232

QY 181 SMSLGLVLYLVHGTMPEFGDHTLVKQISNGAYREPPKPSDAGLIRMLMNPTRRA 240

DB 233 SMSLGLVLYLVHGTMPEFGDHTLVKQISNGAYREPPKPSDAGLIRMLMNPTRRA 292

QY 241 TLEDVASHMWVNW 253

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

XX Homo sapiens.
OS WO2004043361-A2.
PN 27-MAY-2004.
PD 06-NOV-2003; 2003WO-US035268.
PP 08-NOV-2002; 2002US-0425235P.
PR (GENTH) GENENTECH INC.
PX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WJ, Wu TD;
PI WPI; 2004-420067/39.
PX N-PSDB; ADO20171.
DR Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthritis.
XX Claim 7; SEQ ID NO 1080; 1731pp; English.

XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polynuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polynuropathy. This sequence represents a human PRO polypeptide of the
XX invention.

SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;
Best Local Similarity 95.7%; Pred. No. 6.8e-133;

Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGYGKVKARSSGRVLAIKSIRKDKI KDEDLHRRREIIMSSLNHPH 60

DB 53 YEFLETLGKGTGYGKVKARSSGRVLAIKSIRKDKI KDEDLHRRREIIMSSLNHPH 112

QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120

DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172

QY 121 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 180

DB 173 HRDLKLENTLLDANNGIKIADFGLSNLVHKGFLOTFCGSPLYASPEYVNGKPYGPEVD 232

QY 181 SMSLGLVLYLVHGTMPEFGDHTLVKQISNGAYREPPKPSDAGLIRMLMNPTRRA 240

DB 233 SMSLGLVLYLVHGTMPEFGDHTLVKQISNGAYREPPKPSDAGLIRMLMNPTRRA 292

QY 241 TLEDVASHMWVNW 253

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

QY

DB 293 TLEDVASHMWVNW 305

XX Human cancer-associated protein (CAP) HP07-004.
XX Human; cancer-associated protein; CAP; cancer; cytostatic.
XX Homo sapiens.
XX MO2004058146-A2.
XX 15-JUL-2004.
XX 15-DEC-2003; 2003MO-US040081.
XX 17-DEC-2002; 2002US-00322281.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX WPI; 2004-499109/47.
XX N-PSDB; ABD33085.
XX Novel human cancer associated protein encoded within open reading frame
XX of cancer associated gene, useful as targets for diagnosing cancer.
XX Claim 18; SEQ ID NO 26; 182bp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
XX associated (CA) nucleic acids encoding them. The invention also relates
XX to a method for treating cancers involving administering to a patient an
XX inhibitor of CAP, and a method of screening for anticancer activity in a
XX potential drug involving providing a cell that expresses a CA gene,
XX contacting a tissue sample derived from a cancer cell with an anticancer
XX drug candidate and monitoring the effect of the anticancer drug candidate
XX on expression of the CA gene. The CAP proteins are useful for detecting
XX cancer associated with expression of a CAP protein in a test cell sample
XX and for screening for a bioactive agent capable of modulating the
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX cancer, involving determining the expression of a CA nucleic acid in a
XX tissue. This sequence represents a human CAP of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 628 AA:
XX
XX Query Match 96.3%; Score 1293; DB 8; Length 628;
XX Best Local Similarity 95.7%; Pred. No. 6.8e-133;
XX Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 YEFLETLGKGTGKVKKARSSGRVLAISKIRDKIKDEODLHIREIRIEMSLNPHI 60
XX |||||||
XX 53 YEFLETLGKGTGKVKKARSSGRVLAISKIRDKIKDEODLHIREIRIEMSLNPHI 112
XX |||||||
XX 61 IAIHEVFENSSKIVIVMEVYASRGDLVDYISERPRLSERDARHFRQVLSALHYCHONGIV 120
XX ||||||| :||||| :
XX DB 113 IAIHEVFENSSKIVIVMEVYASRGDLVDYISERQOLSREARHFRQVLSAVHCHORRV 172
XX ||||||| :
XX QY 121 HRDLKLENTLLDANGNINKIADFGLSNLHYHGKFLQTFCCGSPLYASPEIIVNGKPYGPEVD 180
XX |||||||
XX DB 173 HRDLKLENTLLDANGNINKIADFGLSNLHYHGKFLQTFCCGSPLYASPEIIVNGKPYGPEVD 232
XX |||||||
XX QY 181 SWSLGVLLYTLVHGTMPEFGQDHKTIVKQISNGAYRPPKPSDACGLIRWLMLVNPTRRA 240
XX |||||||
XX DB 233 SWSLGVLLYTLVHGTMPEFGQDHKTIVKQISNGAYRPPKPSDACGLIRWLMLVNPTRRA 292
XX |||||||
XX QY 241 TLEDVASHMWVNM 253
XX |||||||
XX DB 293 TLEDVASHMWVNM 305
XX |||||||
XX
XX RESULT 14
XX ADJ96620

ID ADJ96620 standard; protein; 672 AA.
XX AC ADJ96620;
XX AC ADJ96620;
XX DT 06-MAY-2004 (first entry)
XX
XX Human calcium/calmodulin-dependent protein kinase α 2 protein SeqID 77.
XX
XX kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
XX PKR; STK; gene therapy; cancer; immune-related disease;
XX cardiovascular disease; brain; neuronal associated disease; metabolic;
XX inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
XX antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;
XX α 2.
XX
XX Homo sapiens.
XX 72.
XX MO2004006838-A2.
XX 22-JAN-2004.
XX
XX 15-JUL-2003; 2003MO-US021730.
XX 15-JUL-2002; 2002US-0395632P.
XX
XX (SUGR-) SUGEN INC.
XX Whyte D, Manning G, Caenepeel S;
XX WPI; 2004-122753/12.
XX N-PSDB; ADJ96554.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
XX preparing a composition for treating diseases or disorders, e.g., cancer,
XX or neurological, immunological or inflammatory disorders.
XX
XX Claim 1; SEQ ID NO 77; 366bp; English.
XX
XX This invention relates to a novel isolated, enriched or purified nucleic
XX acid molecule that encodes a kinase polypeptide. Specifically, it relates
XX to human tyrosine and serine/threonine protein kinases (PK's and STK's),
XX as well as protein kinase-like enzymes. The present invention describes
XX screening methods to identify agonists, antagonists and antibodies that
XX can be used to modulate the activity or function of the mammalian kinase
XX enzymes. As such, these compositions can be used for gene therapy
XX purposes to treat diseases or disorders including cancer, immune-related
XX diseases, cardiovascular disease, brain or neuronal associated disease,
XX metabolic and inflammatory disorders. Accordingly, they exhibit
XX cytostatic, neuroprotective, immunomodulator and antiinflammatory
XX activities. This polypeptide sequence is a human kinase protein sequence
XX of the invention.
XX
XX Sequence 672 AA:
XX
XX Query Match 96.3%; Score 1293; DB 8; Length 672;
XX Best Local Similarity 95.7%; Pred. No. 7.5e-133;
XX Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 YEFLETLGKGTGKVKKARSSGRVLAISKIRDKIKDEODLHIREIRIEMSLNPHI 60
XX |||||||
XX DB 97 YEFLETLGKGTGKVKKARSSGRVLAISKIRDKIKDEODLHIREIRIEMSLNPHI 156
XX |||||||
XX QY 61 IAIHEVFENSSKIVIVMEVYASRGDLVDYISERPRLSERDARHFRQVLSALHYCHONGIV 120
XX ||||||| :||||| :
XX DB 157 IAIHEVFENSSKIVIVMEVYASRGDLVDYISERQOLSREARHFRQVLSAVHCHORRV 216
XX ||||||| :
XX QY 121 HRDLKLENTLLDANGNINKIADFGLSNLHYHGKFLQTFCCGSPLYASPEIIVNGKPYGPEVD 180
XX |||||||
XX DB 217 HRDLKLENTLLDANGNINKIADFGLSNLHYHGKFLQTFCCGSPLYASPEIIVNGKPYGPEVD 276
XX |||||||
XX QY 181 SWSLGVLLYTLVHGTMPEFGQDHKTIVKQISNGAYRPPKPSDACGLIRWLMLVNPTRRA 240
XX |||||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 11:55:01 ; Search time 303.196 Seconds
(without alignments)
10216.077 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atcgagtcgctg9ccttact.....gaatctgctcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgm2_6/prodata/1/ina/5A COMB.seq:*
2: /cgm2_6/prodata/1/ina/5B COMB.seq:*
3: /cgm2_6/prodata/1/ina/6A COMB.seq:*
4: /cgm2_6/prodata/1/ina/6B COMB.seq:*
5: /cgm2_6/prodata/1/ina/6C COMB.seq:*
6: /cgm2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	4 US-09-579-664B-4	Sequence 4, Appl1
2	1893	100.0	2902	4 US-10-355-975A-4	Sequence 4, Appl1
3	1355.4	71.6	2501	4 US-09-799-451-215	Sequence 215, App
4	204.2	10.8	2234	4 US-09-949-016-2384	Sequence 2384, Ap
5	204.2	10.8	2234	4 US-09-949-016-2385	Sequence 2385, Ap
6	204.2	10.8	2234	4 US-09-949-016-1546	Sequence 1546, Ap
7	204.2	10.8	2234	4 US-09-949-016-1547	Sequence 1547, Ap
8	202.6	10.7	2175	4 US-09-984-890-1	Sequence 1, Appl1
9	202.6	10.7	2175	4 US-10-274-194-1	Sequence 1, Appl1
10	189.2	10.0	484	4 US-09-270-767-1257	Sequence 1257, Ap
11	189.2	10.0	484	4 US-09-270-767-16539	Sequence 16539, A
12	186.4	9.8	2908	3 US-09-930-181-1	Sequence 1, Appl1
13	182.4	9.6	2652	3 US-08-557-006C-39	Sequence 39, Appl1
14	182.4	9.6	2761	3 US-08-557-006C-24	Sequence 24, Appl1
15	181.2	9.6	1742	3 US-08-557-006C-38	Sequence 38, Appl1
16	179.8	9.5	2638	2 US-08-677-298-1	Sequence 1, Appl1
17	179.8	9.5	2638	2 US-09-949-016-2363	Sequence 2363, Ap
18	178.2	9.4	2914	4 US-09-949-016-343	Sequence 343, App
19	176.4	9.3	2899	4 US-09-774-528-112	Sequence 112, App
20	174	9.2	3364	3 US-09-930-181-3	Sequence 3, Appl1
21	173	9.1	2337	4 US-10-003-690-3	Sequence 3, Appl1
22	173	9.1	2337	4 US-10-116-326-1	Sequence 1, Appl1
23	173	9.1	2980	4 US-10-003-690-1	Sequence 1, Appl1
24	172.2	9.1	1539	4 US-09-633-328B-1	Sequence 1, Appl1
25	172	9.1	3609	4 US-09-799-875-6	Sequence 6, Appl1
26	172	9.1	5983	4 US-09-799-875-4	Sequence 4, Appl1
27	171.4	9.1	1647	3 US-09-101-146-44	Sequence 44, Appl1

28	165.6	8.7	1229	3 US-09-359-161-4	Sequence 4, Appl1
29	163.6	8.6	1747	3 US-08-557-006C-44	Sequence 44, Appl1
30	161.6	8.5	1863	4 US-09-949-016-4303	Sequence 4303, Ap
31	160.4	8.5	1507	3 US-09-523-849-4	Sequence 4, Appl1
32	160.4	8.5	1544	3 US-09-734-673-1	Sequence 1, Appl1
33	160.4	8.5	2060	3 US-09-523-849-1	Sequence 1, Appl1
34	158	8.3	1736	3 US-08-557-006C-37	Sequence 37, Appl1
35	158	8.3	1783	3 US-08-557-006C-36	Sequence 36, Appl1
36	155.6	8.2	1051	3 US-09-523-849-3	Sequence 3, Appl1
37	154	8.1	2112	4 US-10-116-326-5	Sequence 5, Appl1
38	149.4	7.9	1599	3 US-09-256-465-1	Sequence 1, Appl1
39	149.4	7.9	1599	3 US-09-167-322-3	Sequence 3, Appl1
40	149.4	7.9	1539	4 US-09-023-653-1004	Sequence 1004, Ap
41	147.8	7.8	1254	4 US-09-590-740-3	Sequence 3, Appl1
42	144.4	7.6	2181	4 US-09-417-197-70	Sequence 70, Appl1
43	144.4	7.6	2184	4 US-09-417-197-138	Sequence 138, App
44	144.4	7.6	2610	2 US-09-212-771-1	Sequence 1, Appl1
45	144.4	7.6	2610	3 US-09-091-058-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-579-664B-4
; Sequence 4, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Varca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579, 664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

Query Match	100.0%;	Score 1893;	DB 4;	Length 2902;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1893;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGAGTCGGTGGCTTACTCCAGCGCCGAGCGAGGCTCCGCGCTCCGCGTGGCC	60	
DB	123	ATGAGTCGGTGGCTTACTCCAGCGCCGAGCGAGGCTCCGCGCTCCGCGTGGCC	182	
QY	61	TCGAGAGCGCCCGCGCGCTGGCGGACGGGCTCATGAGTGGCTAACTTGATGAAG	120	
DB	183	TCGAGAGCGCCCGCGCGCTGGCGGACGGGCTCATGAGTGGCTAACTTGATGAAG	242	
QY	121	AACGAGCGGTGAAGCGGCATCAACAACAACAACCTGGGCAACGCTTACGAGTTCTG	180	
DB	243	AACGAGCGGTGAAGCGGCATCAACAACAACAACCTGGGCAACGCTTACGAGTTCTG	302	
QY	181	GAGAGCGTGGGCAAGGCACTTACGGGAAAGGTGAAGAGGACGAGAGAGCTGGGGCGT	240	
DB	303	GAGAGCGTGGGCAAGGCACTTACGGGAAAGGTGAAGAGGACGAGAGAGCTGGGGCGT	362	
QY	241	CTGGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAATGAGAGATCTGCTGCAC	300	
DB	363	CTGGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAATGAGAGATCTGCTGCAC	422	
QY	301	ATACGAGGAGATTTGAGATCATGCTTCACTCAACCAACCCCAATCATTCATTCAT	360	
DB	423	ATACGAGGAGATTTGAGATCATGCTTCACTCAACCAACCCCAATCATTCATTCAT	482	

QY 361 GAAAGTTTGAAATAGACAGCAAGATTGTGATTGTGAGATATGCGACCGAGGGGAT 420
Db 483 GAAAGTTTGAAATAGACAGCAAGATTGTGATTGTGAGATATGCGACCGAGGGGAT 542
QY 421 CTGTATGATTAATCATCACTGAGCGGCGCA CGCTGAGTGA CGCGGAGCGGAGCATTTCTTC 480
Db 543 CTGTATGATTAATCATCACTGAGCGGCGCA CGCTGAGTGA CGCGGAGCGGAGCATTTCTTC 602
QY 481 CGACAGATCGGTGTGCTGCTGCACTACTGCCACAGAA CGGGATGCTTACCGAGATCTC 540
Db 603 CGACAGATCGGTGTGCTGCTGCACTACTGCCACAGAA CGGGATGCTTACCGAGATCTC 662
QY 541 AAGCTGAAAAACATCTTTAGATGCCAATGSAACATCAAGATTGCTGACTTTGGGCTC 600
Db 663 AAGCTGAAAAACATCTTTAGATGCCAATGSAACATCAAGATTGCTGACTTTGGGCTC 722
QY 601 TCCAACTGTACCAACAAGGCAAGTTCTCCAGACGTTCTGTGGGAGCCCTCTTACGCG 660
Db 723 TCCAACTGTACCAACAAGGCAAGTTCTCCAGACGTTCTGTGGGAGCCCTCTTACGCG 782
QY 661 TCGCTGAGATATGTAACGCGAAGCCCTATGTGGGCCCAAGAGTGA CAAGCTGTCTTG 720
Db 783 TCGCTGAGATATGTAACGCGAAGCCCTATGTGGGCCCAAGAGTGA CAAGCTGTCTTG 842
QY 721 GGGGTTCTCTGTACATCTGTGTGATGAGCAACATGCGCTTTGACGGGCGAGATCATAAA 780
Db 843 GGGGTTCTCTGTACATCTGTGTGATGAGCAACATGCGCTTTGACGGGCGAGATCATAAA 902
QY 781 ACACTGTGAAAGCAAAATCAGTAAACGGGGCTTACCGTGAAGCGGCCCAAGCGCTCGATGCC 840
Db 903 ACACTGTGAAAGCAAAATCAGTAAACGGGGCTTACCGTGAAGCGGCCCAAGCGCTCGATGCC 962
QY 841 TGTGGCTGTATCCGGTGGCTGTTATGTGTAAATCCCAACCGGTGGGCGCACTGAGAGAT 900
Db 963 TGTGGCTGTATCCGGTGGCTGTTATGTGTAAATCCCAACCGGTGGGCGCACTGAGAGAT 1022
QY 901 GTAGCCGATGATGGTGGGTCAACTGGGGTTTACCAACGCGGAGTGGGGAAACAGAAAGCC 960
Db 1023 GTAGCCGATGATGGTGGGTCAACTGGGGTTTACCAACGCGGAGTGGGGAAACAGAAAGCC 1082
QY 961 CTGCGTGAAGGAGTGGGCACTCTAGTGTGACTTTTGGCGGGCGCTTCATGCGGAGCTGTGTA 1020
Db 1083 CTGCGTGAAGGAGTGGGCACTCTAGTGTGACTTTTGGCGGGCGCTTCATGCGGAGCTGTGTA 1142
QY 1021 CGTGGCTCTTGGGCGCCCTCTCTGAGAGATGAGCCCAAGGTGTGACGCTTCTTCAAGCAG 1080
Db 1143 CGTGGCTCTTGGGCGCCCTCTCTGAGAGATGAGCCCAAGGTGTGACGCTTCTTCAAGCAG 1202
QY 1081 CAGTGCCTGGAGGTGGAACACTGTACCTGGGTGAGAGGGGCAACTTCTTTAAGAG 1140
Db 1203 CAGTGCCTGGAGGTGGAACACTGTACCTGGGTGAGAGGGGCAACTTCTTTAAGAG 1262
QY 1141 TCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTCT 1200
Db 1263 TCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACCTCT 1322
QY 1201 TCTGCGCTTGGCAAGAGCACTTAACTTCCGAAGGCAATTTCTCAAGAAAAGTCTCT 1260
Db 1323 TCTGCGCTTGGCAAGAGCACTTAACTTCCGAAGGCAATTTCTCAAGAAAAGTCTCT 1382
QY 1261 ACCTCGTCAGGGAGATGATGAGAGGAGACCTCTCAAGAACTCAAGACGGGTGCTGATCTCA 1320
Db 1383 ACCTCGTCAGGGAGATGATGAGAGGAGACCTCTCAAGAACTCAAGACGGGTGCTGATCTCA 1442
QY 1321 GGGCAGACCTGTCCCTGTGTATCCCTGTCTCCAAAGAAAGGCACTCTTAAGAGTCTCA 1380
Db 1443 GGGCAGACCTGTCCCTGTGTATCCCTGTCTCCAAAGAAAGGCACTCTTAAGAGTCTCA 1502
QY 1381 CAGCGTGAATCTGTGTTACTACTCTCTCAAGAGCCGAGAGTCTGGGAAACTTTAGAC 1440
Db 1503 CAGCGTGAATCTGTGTTACTACTCTCTCAAGAGCCGAGAGTCTGGGAAACTTTAGAC 1562
QY 1441 GCCAGTGAATCTGTGTTAGTGGGAGCCCGGTGAGCAAGAGTCTCCACAGGCTTCAAGG 1500

Db 1563 GCCAGTGAATCTGTGTTAGTGGGAGCCCGGTGAGCAAGAGTCTCCACAGGCTTCAAGG 1622
QY 1501 CTCCTCTCCACCGCAAGGCACTTCTCAACTCAATGAGCAAGTCTCCGCAAGCTTCA 1560
Db 1623 CTCCTCTCCACCGCAAGGCACTTCTCAACTCAATGAGCAAGTCTCCGCAAGCTTCA 1682
QY 1561 GAAGGCACTACCCCTTACGACCTTTGGCTCTGAGACCAATGGGCTCTCTCCATCTGCA 1620
Db 1683 GAAGGCACTACCCCTTACGACCTTTGGCTCTGAGACCAATGGGCTCTCTCCATCTGCA 1742
QY 1621 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGAGACGATCTCTGCTCCGAGTCC 1680
Db 1743 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGAGACGATCTCTGCTCCGAGTCC 1802
QY 1681 TTTGACCAATTTGACTTGTGCTGAACTCTCCGAAACCCCACTGAGGGGCTGTGTCT 1740
Db 1803 TTTGACCAATTTGACTTGTGCTGAACTCTCCGAAACCCCACTGAGGGGCTGTGTCT 1862
QY 1741 GTGGAACAACCTGAGGGGCTTGAAGACGCTCTCTCAAGAGTCTGAAGCGATGTGGCAG 1800
Db 1863 GTGGAACAACCTGAGGGGCTTGAAGACGCTCTCTCAAGAGTCTGAAGCGATGTGGCAG 1922
QY 1801 GAATCTGTGGGGATGTAGCTGTTTCTGTGACAGACTGCCAAGAGTGAAGTGAAGCTAC 1860
Db 1923 GAATCTGTGGGGATGTAGCTGTTTCTGTGACAGACTGCCAAGAGTGAAGTGAAGCTAC 1982
QY 1861 AGACAAGCCCTAGAAATCTGCTCAAAAGCTCAGC 1893
Db 1983 AGACAAGCCCTAGAAATCTGCTCAAAAGCTCAGC 2015

RESULT 2
US-10-355-975A-4
; Sequence 4, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355,975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-355-975A-4

Query Match 100.0%; Score 1893; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAATCGGTGCTTACTCTCAAGCGCCGAGAGCGGCTCCCTGGGCTCCGCGCTGGCG 60
Db 123 ATGGAATCGGTGCTTACTCTCAAGCGCCGAGAGCGGCTCCCTGGGCTCCGCGCTGGCG 182
QY 61 TCGGAGAGCGCCCGGCGCTGTGGGAGCGGGCTCATCAAGTGTGCTTAACTGTATGAAG 120
Db 183 TCGGAGAGCGCCCGGCGCTGTGGGAGCGGGCTCATCAAGTGTGCTTAACTGTATGAAG 242
QY 121 AAGCAGCGGTGAAGCGGCACTATCAACAACAACCTCGGCAACGCTTACGAGTTCCTG 180
Db 243 AAGCAGCGGTGAAGCGGCACTATCAACAACAACCTCGGCAACGCTTACGAGTTCCTG 302
QY 181 GAGAGCGTGGGCAAGGCACTTACGAGGAGGTGAAGAGAGCAAGAGAGTCTGGGGCGT 240
Db 303 GAGAGCGTGGGCAAGGCACTTACGAGGAGGTGAAGAGAGCAAGAGAGTCTGGGGCGT 362

QY 241 CTGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAGATGAGAGATCTGTCAC 300
DB 363 CTGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAGATGAGAGATCTGTCAC 422
QY 301 ATACGAGGAGATGAGATCAATGTCCTTCACTCAACCAACCCCATCATTTGCCATCAT 360
DB 423 ATACGAGGAGATGAGATCAATGTCCTTCACTCAACCAACCCCATCATTTGCCATCAT 482
QY 361 GAAGTGTGAGATGAGATCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 420
DB 483 GAAGTGTGAGATGAGATCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 542
QY 421 CTGTATGATTAATCAATGAG 480
DB 543 CTGTATGATTAATCAATGAG 602
QY 481 CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 603 CGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
QY 541 AAGCTGAAAACATCCCTTCTAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
DB 663 AAGCTGAAAACATCCCTTCTAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 722
QY 601 TCCAACTGTACCAAAAGGCAAGTTCTCCAGAGTTCTGTGGAGACCTCTCTACAGCC 660
DB 723 TCCAACTGTACCAAAAGGCAAGTTCTCTCAAGCTTCTGTGGAGACCTCTCTACAGCC 782
QY 661 TGCCTGTAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 720
DB 783 TGCCTGTAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 842
QY 721 GGGGTTCTCTGTATCAATCTGTGTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 780
DB 843 GGGGTTCTCTGTATCAATCTGTGTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 902
QY 781 ACACTGTGAGAGCAAAATCAAGTACAGGGGCTTACCGTAGAGCCGCAAGCGTCCGATGCC 840
DB 903 ACACTGTGAGAGCAAAATCAAGTACAGGGGCTTACCGTAGAGCCGCAAGCGTCCGATGCC 962
QY 841 TGTGGCCTGATCCGGTGGCTGTATATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 900
DB 963 TGTGGCCTGATCCGGTGGCTGTATATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1022
QY 901 GTAGCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1023 GTAGCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 961 CTGTGTGAGGATGAG 1020
DB 1083 CTGTGTGAGGATGAG 1142
QY 1021 CGTGTCTCTCTCGCCCTCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
DB 1143 CGTGTCTCTCTCGCCCTCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1202
QY 1081 CAGGTGCGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1140
DB 1203 CAGGTGCGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1262
QY 1141 TCCGAGAAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1200
DB 1263 TCCGAGAAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1322
QY 1201 TCTGTCCCTGTGCAAG 1260
DB 1323 TCTGTCCCTGTGCAAG 1382
QY 1261 AACTGTGTCAAGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1320
DB 1383 AACTGTGTCAAGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1442

QY 1321 GGGCAGCCTGTCTCTGTCTGTATCTCTGTCTCCAGAGAAAGCATCTTAAAGTCTGCA 1380
DB 1443 GGGCAGCCTGTCTCTGTCTGTATCTCTGTCTCCAGAGAAAGCATCTTAAAGTCTGCA 1502
QY 1381 CAGCGTAATCTGTATCTCTCTGTCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1503 CAGCGTAATCTGTATCTCTCTGTCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
QY 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1563 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
QY 1501 CTCTCTCTCCAG 1560
DB 1623 CTCTCTCTCCAG 1682
QY 1561 GAAGGACATACCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1683 GAAGGACATACCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742
QY 1621 GCCGAGCCAG 1680
DB 1743 GCCGAGCCAG 1802
QY 1681 TTGACCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1803 TTGACCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
QY 1741 GTGAGACATCTGAG 1800
DB 1863 GTGAGACATCTGAG 1922
QY 1801 GAATCTGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1923 GAATCTGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
QY 1861 AGACAAGCCCTAGAGATCTGCTCAAGCTCAGC 1893
DB 1983 AGACAAGCCCTAGAGATCTGCTCAAGCTCAGC 2015

RESULT 3
US-09-799-451-215
Sequence 215, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, RyLe
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feijian
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Mehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Dimahe, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 215
LENGTH: 2501

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106) .. (1989)
; US-09-799-451-215

```

Query Match	71.6%	Score 1355.4	DB 4	Length 2501
Best Local Similarity	84.0%	Pred. No. 0		
Matches 1562; Conservative	0	Mismatches 276	Indels 21	Gaps 2

QY	47	CTCTCCGCGCTGCGCTCGGAGAGCCCGCGCGCTGCGGCGAGCGGGCTCATCAAGTGCCTTA	106
Db	140	CAACTCCCTCGCGCGGAGAGCTACCGCGCGCTGCGGAGGGGCTGATCAAGTGCCTCA	199
QY	107	AACCTTGATGAGAGAGGCGGGTGAAGCGGACCATCAACACAACTCGCGGACCC	166
Db	200	AGCCCTTAATGAAAGAGAGCGGGTGAAGCGGACCAACAGACAACTCGCGGACCC	259
QY	167	GCTACGAGTTCCTGAGAGCGCTGGGCGAAGGGGACCTACGGGAAGTGAAGAAAGGACGAG	226
Db	260	GCTACGAGTTCCTGAGAGCCTGGGCGAAGGGGACCTACGGGAAGTGAAGAAAGGCGCGG	319
QY	227	AGAGCTCGGGGCGTCTGGTGGCCATCAAGTCATCAGGAAAGACAAATCAAGATGAGC	286
Db	320	AGAGCTCGGGGCGCTGGTGGCCATCAAGTCATCAGGAAAGACAAATCAAGATGAGC	379
QY	287	AGAGTCTGCTGCAATACGAGGGGAGATGAGATCATATCTTCACTACACACCCGACCA	346
Db	380	AAGATCTGATGCAATACGAGGGGAGATGAGATCATATCTCACTACACACCCCTCA	439
QY	347	TCAATTGCATTCATGAAGTGTGTTAGAAATAGCAGCAAGATTGTGATTTGATGAGATAG	406
Db	440	TCATTTGCCATTCATGAAGTGTGTTAGAAAGCAGCAGATGTGTATGTCATGAGATAG	499
QY	407	CCAGCCGAGCGGATCTGTATGATTAATCAATGATGAGCGGCGACGGCTGATGAGCGGAGC	466
Db	500	CCAGCCGGGCGGACCTTATATCAATCAACGAGCGGCGACGACTGATGAGCGGAGG	559
QY	467	CCAGGCAATTTCTTCGAGAGATCGTGTCTGCGCTGCACCTACTGCGACCGAAGCGGATCG	526
Db	560	CTAGGCAATTTCTTCGAGAGATCGTGTCTGCGCTGCACCTATTGGCATTAGAACGAGATTG	619
QY	527	TTCAACCGAGATCTCAAGCTGGAAAAACATCCTTCAAGATGCGCAATGAGAAACATCAAGATTG	586
Db	620	TTCAACCGAGATCTCAAGCTGGAAACATCCTTCAAGATGCGCAATGAGAAATTCAGAGATTG	679
QY	587	CTGACTTTGGCTCTCCAACTGTACCAAAAGCGAAGTTCTTCAAGCGTTCTGTGAGGA	646
Db	680	CTGACTTTCGTCTCTCCAACTCTACCATCAAGGCAAGTTCTTCAAGCATTTCTGTGAGGA	739
QY	647	GCCCTCTCTAGCGCTGCGCTGAGATGTCAGCGGAAAGCCCTATGTGGGGCCCAAGGTGG	706
Db	740	GCCCCCTCTATGCTTCGCCAGAGATTTGTCATATGGGAAGCCCTTACACAGGCCCAAGAGGTGG	799
QY	707	ACAAGCTGATCTTGAGGCGTTCTCCTGTACATCTGTGTGCATGTGGACATGACCTCTTGACG	766
Db	800	ACAAGCTGATCTTGAGGTGTTCTCCTGTACATCTGTGTGCATGTGGACATGACCTTTTATG	859
QY	767	GGCAGGATCTATAAAACTGTGTGAAGCAATCACTAACGGGGCTTTACCTGTGAGCGGCCCA	826
Db	860	GGCAGTGCATTAATATCTATGTGAACAGATCAACAAACGGGGCTTACCGGGAGCCACTTA	919
QY	827	AGCGGTCCAGATGCGCTGTGGCCTGATCCGGGTGGCTTTAATGTGTAAACCCACCGGTGGG	886
Db	920	AACCTCTGATGCTGTGTGGCCTGATCCGGGTGGCTTTTGTATGTGTAAACCCACCGCGGG	979
QY	887	CCAACATGAGAGATGTAGCAAGCATTTGTGTGGTCAACTGGGGTTTACACACCGGAGTCG	946
Db	980	CCAACCTGAGAGATGTGGCAGTCACTGTGTGGTCAACTGGGGCTTACCGGACCGGAGTGG	1039
QY	947	GGGAAACAGGAAGCCCTGCGTGAAGGTGGGACCTTAATGTGTGACTTTTGGCCGGGCTTCA	1006

Ds	1040	GAGAGCAGGAGGCTCCGATAGAGGTGGGACACCTGGCAGTGACTCTGCCCCGGCCTCCA	1099
Qy	1007	TGGCGGACTGTTAAGTGTCTCTCGCGCCCCCTCTCTGAGAAATGAGCCAGGTGTGA	1066
Ds	1100	TGGCTGACTGGCTCCGGGGTTCTCTCCGCCCTCTCTGAGAAATGGGGCCAAAGGTGTGA	1159
Qy	1067	GCCTCTTCAAGAGCAGTGGCCGGGAGGTGGAACCTGATCTTGGGGCTGAGAGCGGCAAC	1126
Ds	1160	GCCTCTTCAAGAGCAGTGAACCTGATGGGGGAGACACCCCTGGGCTTGAAGCGCAGC	1219
Qy	1127	ATTCTCTTAAGAAAGTCCCGAAAGGAAATGACATAGGCTCAAAATCTCTCAAGGTGACCCGG	1186
Ds	1220	ATTCCGTCAAGAAAGTCCCGCAGAAGAGATGACATAGGCCAGTCTCTCAAGTGAACACGG	1279
Qy	1187	CTGAGGAAATCCTTCTCTGCCCCCTGGCAAGACAGCTTAAAGCTTCGAAAGGCAATCTCA	1246
Ds	1280	CTGATGACACTGCCCATGGCCCTGGCAGAGACCACTCAAGCTGCCAAAGGGCAATCTCA	1339
Qy	1247	AGAAAAAGTCCCTTAACCTGTCTACAGGGAGGTACAGAGAGAACCTTCAGAAACTCAGACCGG	1306
Ds	1340	AGAAAGAGGTGTCAACCTCTGTCAAGAGGGGTACAGAGAGAACCTTCGGAACCTCAGGCCAA	1399
Qy	1307	TGCGTGAATACTCAAGGAGCAGCCTGTCCTGTGATCCCTGCTCCCAAGGAAAGGCATCC	1366
Ds	1400	TCCCTGACAGCCCAAGGAGGCTGCC-----CCCTGCTCCCAAGAGGCAATTC	1450
Qy	1367	TTAAGAAATCTGACAGCGTGAATCTGTGATCTATCTCTCTCCAGAGCCGACGAGTCTG	1426
Ds	1451	TCAGAAAGCCCCGACAGCGCAGTCTGGCTACTACTCTCTCCGAGCCCAATGAATCTG	1510
Qy	1427	GGGAATCTTGAAGCCCAAGTATGTGTGTGAAATGGGGAGCCCGTGGACACAAAGTCTC	1486
Ds	1511	GGAGTCTTGGACCCAGGACCAAGCAAGTGTGTGAAATGGGGAAATCCCAAGAGCAAGAGCTTC	1570
Qy	1487	CACAGGCTTCAGGGGCTCCTCCATCCACGCAAGGSCATTCTCAAACTCAATGGCAAGTTCT	1546
Ds	1571	CGGAAGCTTCAGGGCTGCTCTCTCAATGGCAAGGCAATCTCAAACTCAATGGCAAGTTCT	1630
Qy	1547	CCCGCAGACCTTGAAGAGCACTACCCCTAGACCTTTGGCTCCCTGAGCAAACTGGCCT	1606
Ds	1631	CCAGACAGCCTTGGAGCTCGGGGCCCCACCACTTCGGGTCCCTGAAATGAATCTGGCC	1690
Qy	1607	CCTCCATCTCTGACAGCCCGGCCACAGCGGCCCTTGAAGGGCTGTGAGTGAAGACAGATCC	1666
Ds	1691	CACCTCGCCCCCTGGCCCGGGCCAGCGGACCTCTCAGGGGGCTGTGAGGAGAGACAGATCC	1750
Qy	1667	TGTCCTCCGAGTCCCTTGAACCAATGGAATCTTGCTGCTGAAGCTTCCCGAAACCCCACTGA	1726
Ds	1751	TGTCCTCTGAGTCTCTTTAACAAGCTGGAATCTTGCTGAACGGCTTCCAGAGCCCCCACTGC	1810
Qy	1727	GGGGCTGTGTGTCTGTGACAACCTGAGGGGGCTTTGAGCAGCCTCTCCAGAA-----	1780
Ds	1811	GGGGGCTGTGTGTCTGTGACAACCTCAGGGGGCTTTGAGAGGCCCCCTCAGAGGGCCCTG	1870
Qy	1781	-----GTCTGAAGCGATGTGTGAGGAAATCTTGGGGGATAGCTGCTTTTCTCTGACAG	1834
Ds	1871	GAACTGCTGTGGGGGCTGGCGGAGGAAATCTTTGGGGGAGCACTGCTTTTCCCTGACAG	1930
Qy	1835	ACTGCCAAGAGGTGACTGACAGCTTACAGACAAAGCCCTAAGAAATCTGTCAAAAGCTACG	1893
Ds	1931	ACTGCCAAGAGGTGACAGCACTTACGGAAGGCACTGAGGGGTCTGCTCAAAAGCTACCC	1989
RESULT 4			
US-09-949-016-2384			
: Sequence 2384, Application US/09949016			
: Patient No. 681239			
: GENERAL INFORMATION:			
: APPLICANT: VENTER, J. Craig et al.			
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
: FILE REFERENCE: CL001037			
: CURRENT APPLICATION NUMBER: US/09/949, 016			

```

? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 2384
? LENGTH: 2224
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-2384

```

Query Match	10.8%;	Score 204.2;	DB 4;	Length 2224;
Best Local Similarity	55.0%;	Pred. No. 1.2e-47;		
Matches 423; Conservative	0;	Mismatches 343;	Indels 3;	Gaps 1;

QY	168	CTACGAGTCTCTGGAGACGCTGGGCGAAGGGGACCTACGGGAAGGGAAGGACGAGGA	227
Db	57	CTACCGGCTCTCAAGACCAATTGGCAGAGGTATATTTGGCAAGGTGAAGTTGGCCGACGA	116
QY	228	GAGCTCGGGGCGTCTGTGGCCATCAAGTCATCAGGAAGACAAATCAAAATGAGCA	287
Db	117	CATCTCGACTGGGAAAGAGGTAGCTGTGAMGATCATGACAAAGCTCAACTGAATCTCTC	176
QY	288	GGATCTCTGCACATACGGAGGGAGATTGAATCATGTCTTCACTCAACCAACCCCAAT	347
Db	177	CAGCTCCAGAAACTATTTCCGCGAAGTAAAGATATGAAAGTTTGAATCATCCCAAT	236
QY	348	CATTCGCATCATGAGAGTGTGGAATPAGCAGCAAGATGTGATGTGCATGAGATAGC	407
Db	237	AGTTAAATTAATTTGAATGATTTGAGACTGAGAAAACGCTCTACCTGTTCATGAGATACG	296
QY	408	CAGCGAGCGCATCTGTATGATTAATCATCAGTGAAGCGGCCACACGCTGATGAGCGGACGC	467
Db	297	TAGTGGGCGAAGGATATTGATTAACTAGTGGCTCATGGCAGGATGAAAGAAAAGAGGC	356
QY	468	CAGGCATTTCTTCCGACAGATCGTGTCTGCCCTGCACTACCTGCCACCAAGACGGATGT	527
Db	357	TCGAGCCAAATTCGCGCAGATAGTGTCTGTGAGATACGTCAACCGAAGTTATATGT	416
QY	528	TCACCGAGATCTCAAGCTGGAAGAAACATCCTTCTAGATGCAATGGAACATCAAGATTGC	587
Db	417	CCATAGAGACTTAAAGCAGAAAACCTGCTCTTGATGATCTGATATGAACATCAAGATTGC	476
QY	588	TGACTTTGGCCTCTCCAACCTGTACCAAAAGCGAAGTTCTCTCCAGAGCTTCTGTGGAG	647
Db	477	AGACTTTGGCTTCAAGCAATGAATTCACCTTTGGGAAACAAAGCTGACACCTTCTGTGGAG	536
QY	648	CCCTCTCTACGCTTCGCTGAGATAGTCAACGGGAAGCCATATGTGGGCCCAAGGTGGA	707
Db	537	TCCCCCTTATCTGCTGCCCAAGAACTCTTCAAGGCAAAAATATGATGGAACCCGAGGTGA	596
QY	708	CAGCTGTGCTTGGGCGCTTCTCTGTGACATCTGATGCAATGCAACATAGCCCTTTGACGG	767
Db	597	TGTGTGAGGCTTAGAGATTAATCTCTATACTGGTACGGGATCCCTGCTTTTGAATGG	656
QY	768	GCAGAGATCATTAACACTGTGGAAGCAAAATCAGTAAACGGGGCTTACCGTGAAGCGGCCCA	827
Db	657	ACAGAACTCAAGAGAGCTGCGGGAAACGGGTACTGAGGGGAAAAATATCCGATATTCATTTCTA	716
QY	828	GCCGTCCGATGCTGTGTG--GCTGATCCGGTGCCTGTAAATGTGGAACCCCAACCCGTG	884
Db	717	CATGTCCACGGACTGTGAAAACCTGCTTAAAAAATTTCTCAATTTCTTAATCCCAAGAGAG	776
QY	885	GGCACTAGAGAGATGAGCAATGATCAATGATGTGTGCTCAACTGGGGTTAC	933
Db	777	AGGCACTTAAAGCAATCATGAAAGATCGATGATGATGTGGGTAC	825

```

US-09-949-016-2385
; Sequence 2385, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2385

```

Query Match	10.8%	Score 204.2;	DB 4;	Length 2224;
Best Local Similarity	55.0%	Pred. No. 1.2e-47;		
Matches 423; Conservative	0;	Mismatches 343;	Indels 3;	Gaps 1;

OY	168	CTAAGAGTCTCTGGAGACGCTGGGCAAGGGGCAAGCTTACCGGAAAGGTGAAAGAGCGACGA	227
Db	57	CTACCGGCTCTCAAGACCAATTGGCAAGGGTAATTTGGCAAGGTGAAGTTGGCCCGACA	116
OY	228	GAGCTCGGGGCGTCTGGTGGCCATCAAGTCATCAGGAAAGACAAATCAAAAGATGACA	287
Db	117	CATCTCTAGCTGGGAAAAGGTAGCTGTGAAGATCATTGACAAGACTCAACTGAATCTCTC	176
OY	288	GGATCTGTGCACTAAGAGAGGAGATTGAATCATGTCTTCACTCAACCAACCCCAAT	347
Db	177	CAGCTTCAGAAACTTATTCGCGAAGTAAGATTAATGAAGTTTGAATCATCCCAAT	236
OY	348	CATTGCATTCATGAAGTGTTTGAATATGACAGAAAGTTGTGATTTGTCAATGAGTATGC	407
Db	237	AGTTAAATTAATTTGAAGTATTGACCTGAGAAAACGCTCACTTGTCAATGAGTACGC	296
OY	408	CAGCCGAGCGCATCTGTATGATTTATCATCAAGTGAACGGCCACAGGCGTGAAGCGGAGCG	467
Db	297	TATGGGCGGAGGATATTGTAATTCATGAGCTTCATGGCAGGATGAAGAAAAGAGGC	356
OY	468	CAGGCATTTCTCCGACAGATCGTGTGCGCCCTGCACTACAGCCACCAACGGAGTGT	527
Db	357	TCGAGCCAAATTTCCGCGAGATAGTGTCTGCTGTGCAGTACGTGACCAAGAGTTATGT	416
OY	528	TCACCGAGATCTAAGCTGAAAAATCTCTTCTAGATGCCAATGAGAAAATCAAGATTGC	587
Db	417	CCATAGAGACTTAAAGGACAGAAACCTGCTCTTGAGTGCATATGAACATCAAGATTGC	476
OY	588	TGACTTTTGGCTCTCCAAACCTGTACCAAAAGGAATTCCTCCAGAGCTTCTATGGAG	647
Db	477	AGACTTTTGGCTTACGACATGATTAATCACTTTTGGAAACAACTGGAACACTTCTGTGGAG	536
OY	648	CCCTCTCTACGCTCGGCTGAGATAGTCAACGGGAAGCCATATGTGGGCCAGAGTGA	707
Db	537	TCCCTCTTAATGCTGCCCAAGAACTCTTCAGGGGCAAAAAATATATATGAGCCGAGGTGA	596
OY	708	CAGCTGGTCTCTGGGCGTTCTCTCTGTACATCTTGGTGCATGGCACATGCTTTGACGG	767
Db	597	TGTGTGAGAGCTTAAGAGATTATCTCTATPACCTGTCACTGTCAAGCGGATCTCGCTTTTGAATGG	656
OY	768	GCAGAGTCATTAACAACCTGTGAGCAAAATCAGTAAAGGGCTTACCGTGAAGCGCCACA	827
Db	657	ACAAACCTCAAGGAGGTGGGGGAACGGGTATCTGAGGGGAAAAATATCGTATTAATTCATA	716
OY	828	GCGGTCCGATGCTGTG---GCTGATTCGGGTGGCTGTTAATGTTGAACCCACCCGCTGC	884

Db 717 CATGTCCAGGAGCTGTGAAAACCTGCTTAGAATAATTCTATTCTTATCCACAGAG 776
Qy 885 GGCCACACTGGAGATGAGCACTCATGTGGGTCACTGGGTAC 933
Db 777 AGGCATCTTAGAGCAATCATGAAGAATCATGATGATGATGGGTAC 825

RESULT 6

US-09-949-016-1546
; Sequence 1547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1546
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1546

Query Match 10.8%; Score 204.2; DB 4; Length 2950;
Best Local Similarity 55.0%; Pred. No. 1.4e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

Qy 168 CTACGAGTTCCTGGAGCGCTGGCAAGGCGACCTAGGGAAGGTGAAGAGCAGAGA 227
Db 464 CTACCGGCTCTCAAGACCAATGGCAAGGGTAATTTGGCAAGGTGAAGTGGCCGACA 523
Qy 228 GAGCTCGGGCGCTGTGGCCATCAATCCATCAGAAAGACAAATCAAGATGAGCA 287
Db 524 CATCTGACTGGAGAAAGAGTAGCTGGAAGATCATTTGACAGCTAAGTCTCTC 583
Qy 288 GATCTGCTGCACATACGAGAGGAGATGATCATCTTCACTCAACCACTCCCAT 347
Db 584 CAGCTTCAGAAACTATTCGCGAAGTAAGAAATTAAGAGTTTGAATCATCCAAAT 643
Qy 348 CATTCGATCATGAAGTGTGGAATAGCAGCAAGATTTGTATGATGAGATGTC 407
Db 644 AGTTAAATTTTGAAGTGAATGAGCTGAGCAAAACCTCTCACTTCTCATGAGTACG 703
Qy 408 CAGCCGAGGAGATCTGATGATTAATCATGAGCGGCAAGGCTGAGAGGCGAGC 467
Db 704 TAGTGCGAGAGAGTATTTATTAATCTAGTGCTCATGAGAGAGTGAAGAAAAAGGC 763
Qy 468 CAGGCAATTTCTCGACAGATGCTGTGCTGCTGCACTACTGCGCAGCAGAGCGATCGT 527
Db 764 TCAGGCAAAATTCGCGAAGTAAAGTGTGCTGCTGCACTACTGCAAGAAAGTTATGT 823
Qy 528 TCACGAGATCTCAAGTGAAGAAACATCTTCTAGATGCCAATGAAACATCAAGTTGC 587
Db 824 CCAATAGAGCTTAAGGCAAGAAACCTGCTTGGATGCTGATTAATCAATCAAGATTGC 883
Qy 588 TGAATTTGGCTCTCCAACTGTAACCAAGCAAGCAATTCCTCCAGAGCTTTGGGAG 647
Db 884 AGACTTTGGCTTCAAGCAATGAATTTCACTTTGGAGCAAGCTGAGCACTTCTGTGAG 943
Qy 648 CCTCTCTACGCTCGCTGATGATGATCAAGGAGACCTATGTTGGGCCAGAGTGA 707
Db 944 TCCCCCTTATGCTGCCCAAGACTCTTCCAGGGGAAAAAATATGATGAGACCCAGAGTGA 1003

Qy 708 CAGTGTCTCTGGGGCTTCTCTGTATATCTGTGATGAGCAAGCAATGCCCTTGAAG 767
Db 1004 TGTGTGAGCTTAGAGATTAATCTTATACATGAGTACGAGATCCCTGCTTTGATGG 1063
Qy 768 GCAGATCATATAACAATCTGTGAAGCAAAATCAATTAACGGGGCTTACCGTGAAGCCCA 827
Db 1064 ACAGAACTCAAGAGAGTGGGGAGACGGGTACTGAGGGGAAAAATACGTTATTCATTTCTA 1123
Qy 828 GCCGTCCGATCCTGTG--GCCGTATCCGCTGCTGTTAATGATGAAGCAAGCCGCTG 884
Db 1124 CATGTCCAGGAGCTGTAAGAAACCTGCTTAAGAAATTTCTATTCATTCACAGAGAG 1183
Qy 885 GGCCACACTGAGAGATGAGCACTCATGTGGGTCACTGGGTAC 933
Db 1184 AGGCATCTTAGAGCAATCATGAAGAATCATGATGATGATGGGTAC 1232

RESULT 7

US-09-949-016-1547
; Sequence 1547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1547
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1547

Query Match 10.8%; Score 204.2; DB 4; Length 2950;
Best Local Similarity 55.0%; Pred. No. 1.4e-47;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

Qy 168 CTACGAGTTCCTGGAGCGCTGGCAAGGCGACCTAGGGAAGGTGAAGAGCAGAGA 227
Db 464 CTACCGGCTCTCAAGACCAATGGCAAGGGTAATTTGGCAAGGTGAAGTGGCCGACA 523
Qy 228 GAGCTCGGGCGCTGTGGCCATCAATCCATCAGAAAGACAAATCAAGATGAGCA 287
Db 524 CATCTGACTGGAGAAAGAGTAGCTGGAAGATCATTTGACAGACTGAATCTCTC 583
Qy 288 GATCTGCTGCACATACGAGAGGAGATGATCATGCTTCACTCAACCACTCCCAT 347
Db 584 CAGCTTCAGAAACTATTCGCGAAGTAAGAAATTAAGAGTTTGAATCATCCAAAT 643
Qy 348 CATTCGATCATGAAGTGTGGAATAGCAGCAAGATTTGATGATGAGAGATGTC 407
Db 644 AGTTAAATTTTGAAGTGAATGAGCTGAGCAAAACGCTTACCTTGTCAATGAGTACG 703
Qy 408 CAGCCGAGGAGATCTGATGATTAATCATGAGCGGCAAGGCTGAGTGAAGCGGAGC 467
Db 704 TAGTGCGAGAGAGTATTTGATTAATCTAGTGCTCATGAGAGATGAAGAAAAAGAGGC 763
Qy 468 CAGGCAATTTCTCGACAGATGCTGTGCTGCTGCACTACTGCGACCAAGAGGATCGT 527
Db 764 TCAGGCAAAATTCGCGAAGTAAAGTGTGCTGCTGCACTGCTGCAAGAAAGTTATTTG 823
Qy 528 TCACGAGATCTCAAGTGAAGAAACATCTTCTAGATGCAATGAAACATCAAGTTGC 587
Db 824 CCAATAGAGCTTAAGGCAAGAAACCTGCTTGGATGCTGATGATGAGATCAAGATTGC 883

QY 588 TGAATTGGCTCTCCAACTGTACCAAGGCAAGTCTCTCAAGCGTTCTGTGGAG 647
DB 884 AGACTTTGGCTTCAAGCAATGAATTCCTTTGGGAAACAAGTGGACACTTTCTGTGGAG 943
QY 648 CCTCTCTAGAGCTCGCTGAGATGATGACAGGAAACCCCTATGTGGGCCCCAGAGTGA 707
DB 944 TCCCTCTTAGTGTGCCCCCAAACTCTTCAAGGGCAAAAATATGATGAGCCCGAGGTGA 1003
QY 708 CAGCTGTCTCTGGGCTTCTCTGTATCTGTGTGATGAGCAATGCGCTTTGAAG 767
DB 1004 TGTGTGAGGCTTGAAGATTAATCTCTATACATGTGTGAGGAGATCCTGCTTTTATGG 1063
QY 768 GCAAGATCATATAACATGTGTGAAGCAATCAATGAGGAGCTTACCGTGAGCGCCAA 827
DB 1064 ACAGAACTCAAGAGAGCTGGGGAACGGGTACTGAGGGGAAATACCGTATTCATTCTA 1123
QY 828 GCGGTCCGATGCTGTG--GCTGTATCCGCTGTATGTGTGAACCCCAACCGTGC 884
DB 1124 CATGTCCACGAGCTGTGAATAAAGCTTGAAGAAATTTCTCATTTCTTAATCCAGCAAG 1183
QY 885 GGCACACATGAGATGTAGCAGATCATGTGTGGGTCAACTGGGGTTAC 933
DB 1184 AGGCACTTTAGAGCAATCATGAAGATCGATGATGTGGGTAC 1232

RESULT 8

US-09-984-890-1
; Sequence 1, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-890-1

Query Match 10.7%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 3,4e-47;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 168 CTACGAGTTCTGTGAGACGCTGGGCAAGGCACTCAAGGAGGTGAAGAGCAAGAGA 227
DB 156 CTACCGGCTCTCAAGCAATGTGGCAAGGTTAATTTTCCAAAGGTGAAGTTGGCCGACA 215
QY 228 GAGCTGGGGGCTGTGTGGCCATCAATCATCAGGAAAGCAAAATCAAGATGAGA 287
DB 216 CATCTGACTGAGGAAAGAGTACTGTGAATCATTTGACAAAGCTCACTGAACTCTC 275
QY 288 GATCTGTGCAATACGAGGAGATGATGATCATCTTCACTCAACACCCCAAT 347
DB 276 CAGCTTCCAGAAATCTATTCGCGAAGTAAAGATTAATAGATTTTGAATCATCCCAAT 335
QY 348 CATTCGATCCATGAAGTGTGAGAAATAGCAAGAAATTTGTATGTCAATGAGATGAC 407
DB 336 AGTTAAATTTATTTGAATGATGAGTGAAGAAACCTCTACCTGTCAATGAGATGAC 395
QY 408 CAGCCGAGGAGATCTGTATGATTAATCATCACTGAGCGGCAACGGCTGAGTGAAGCGGAGC 467
DB 396 TAGTGGGAGAGATATTGATTAATCAATGAGCTCATGCGAGATGAAAGAAAGAGGC 455
QY 468 CAGGCAATTTTCTGAGAGATGTGTGTGCTGCTGCACTACTGCAACAGAGGAGTGT 527
DB 456 TCAGGCCAAATTCGCGAGGATGATGTCTGTGTGCAATGATGATGATGATGATGATGAT 515

QY 528 TACCGAGATCTCAAGCTGGAAGAAATCTCTTCTAGATGCCAATGGAAGAAATCAAGATTC 587
DB 516 CCAATAGAGACTTAAGGCAAGAAACCTGCTCTTGATGTGATGATGAATCAATGAAGATTC 575
QY 588 TGAATTGGCTCTCCAACTGTACCAAGGCAAGTCTCTCAAGCGTTCTGTGGAG 647
DB 576 AGACTTTGGCTTCAAGCAATGAATTCCTTTGGGAAACAAGTGGACACTTTCTGTGGAG 635
QY 648 CCTCTCTAGAGCTCGCTGAGATGATGACAGGAAACCCCTATGTGGGCCCCAGAGTGA 707
DB 636 TCCCTCTTAGTGTGCCCCCAAACTCTTCAAGGGCAAAAATATGATGAGATCCAGAGTGA 695
QY 708 CAGCTGTCTCTGGGCTTCTCTGTATCATCTGTGTGATGAGCAACATGCGCTTTGAAG 767
DB 696 TGTGTGAGGCTTGAAGATTAATCTCTATACATGTGTGAGGAGATTCCTGCTTTTATGG 755
QY 768 GCAAGATCATATAACATGTGTGAAGCAATCAATGAGGAGCTTACCGTGAGCGCCCA 827
DB 756 ACAGAACTCAAGAGAGCTGGGGAACGGGTACTGAGGGGAAATACCGTATTCATTCTA 815
QY 828 GCGGTCCGATGCTGTG--GCTGTATCCGCTGTATGTGTGAACCCCAACCGTGC 884
DB 816 CATGTCCACGAGCTGTGAATAAAGCTTGAAGAAATTTCTCATTTCTTAATCCAGCAAG 875
QY 885 GGCACACATGAGATGTAGCAGATCATGTGTGGGTCAACTGGGGTTAC 933
DB 876 AGGCACTTTAGAGCAATCATGAAGATCGATGATGTGGGTAC 924

RESULT 9

US-10-274-194-1
; Sequence 1, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-274-194-1

Query Match 10.7%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 3,4e-47;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 168 CTACGAGTTCTGTGAGACGCTGGGCAAGGCACTCAAGGAGGTGAAGAGCAAGAGA 227
DB 156 CTACCGGCTCTCAAGCAATGTGGCAAGGTTAATTTTCCAAAGGTGAAGTTGGCCGACA 215
QY 228 GAGCTGGGGGCTGTGTGGCCATCAATCATCAGGAAAGCAAAATCAAGATGAGA 287
DB 216 CATCTGACTGAGGAAAGAGTACTGTGAAGATCATTTGACAAAGCTCACTGAACTCTC 275
QY 288 GATCTGTGCAATACGAGGAGATGATGATCATCTTCACTCAACACCCCAAT 347
DB 276 CAGCTTCCAGAAATCTATTCGCGAAGTAAAGATTAATAGATTTTGAATCATCCCAAT 335
QY 348 CATTCGATCCATGAAGTGTGAGAAATAGCAAGAAATTTGTATGTCAATGAGATGAC 407
DB 336 AGTTAAATTTATTTGAATGATGAGTGAAGAAACCTCTACCTGTCAATGAGATGAC 395
QY 408 CAGCCGAGGAGATCTGTATGATTAATCATCACTGAGCGGCAACGGCTGAGTGAAGCGGAGC 467
DB 396 TAGTGGGAGAGATATTGATTAATCAATGAGCTCATGCGAGATGAAAGAAAGAGGC 455

Db	48	ACTCTTAATAATGATGTCAGATGATGTAATTTCTTAAGAACTAGCTGTGTGATATGGCAAAATTA	543
Qy	657	CGCCTCGCCCTGAATATGATCAAGGGAAAGCCCTATATGTGGCCCAAGAGTGGACAGCTGTGC	718
Db	544	TGACGACACCGGAAGTATCTCAGAAAGGCTGTATGCGGGTCTTGAGTTGATATCTGGAG	603
Qy	717	TCTGGGCGTTCTCTGTATCATCTGTGTGCATGGACATAGCCCTTTTGAACGGGACAGATCA	776
Db	604	CTGGGAGTTATCTGTATGCCCTTCTGTGGACCCCTCCCGTTTGACGATGAGCAAGT	665
Qy	777	TAAACAACGTGTAAACAATCAATACAGGGGCTTACCTAGCCGCCCAAGCCGTCCGA	835
Db	664	GCCTAGCCTTTTAAAGAGATCCGAGGGGTGTGTTCATCATCCGAGATATTCAAACG	723
Qy	837	TGCCTGTGGC---CTGATCCGGTGGCTGTAAATGTGACACCCACCCGTGGGACACACT	893
Db	724	TTCTATTCGCACCTGTGCTATGCAATGTGTGACGGGTGACCCCTTGAAGCAGACACTAT	783
Qy	894	GGAGAGATGACCAAGTCATTTGGTGGGTCAACTGGGGTTACACCA	937
Db	784	CAAAAGCATTCGAGACCATGAATGGTTTAAACAGGATTTGGCCCA	827

```

RESULT 14
US-08-557-006C-24
Sequence 24, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Beri, Rajinder K.
APPLICANT: Carling, David
APPLICANT: Forder, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PMH37588/UST
CURRENT APPLICATION NUMBER: US/08/557, 006C
CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 2761
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat
OTHER INFORMATION: liver AMP protein kinase
US-08-557-006C-24

```

Query Match	9.6%	Score 182.4	DB 3	Length 2761		
Best Local Similarity	53.0%	Pred. No. 2.2e-41				
Matches 4377	Conservative 0	Mismatches 381	Indels 6	Gaps 2		
QY	120	GAAGCAGCGCGGTGAAGCCGCA	CCATCA	CAAAACAACTGCGGCA	CCGCTACGAGTTCTT	179
Db	20	GAAATGGCTGAGAAGCAGAAAGCA	CGCGGCTGTGAAGATCGACA	CTACGTCTCGG		79
QY	180	GGACACGCTGGGCAAGGCGACCTCT	CGGAAAGTGAAGAAAGCAGAGAGCTCG	--GG		236
Db	80	GGACACCTGGGCGCTGGCACCTT	TGGCAAGTGAAGTTGGAGAACATCAATTACAGG			139
QY	237	GCGTCTGGTGCCATCAAGTCATCAG	AAAGACAAATTCAAAGATGACAGATCTCT			296
Db	140	CCATPAAAGTCAGATTAAATCTT	AAATAGACAGAAAGATTCGCA	GTTAGATGTTGG		199
QY	297	GCATATACGAGGGGAGATTGAGAT	ATATCTTCACTCAACACCCCAATCATTTGCCAT			356
Db	200	AAAATTAAGACGAATTCAAATCTT	AAACTCTTGTGTAATCTCTCAATATTAATCACT			259
QY	357	CCATGAAGTGTTTGAAGATAGCAG	ACAGATTGTGATTCATGAGATATGCCAGCCGAG			416

Db 260 CTACCAAGTATCAGCACTCCACAAGACTTTTATATGTAATGAATATATGTCTTGAGG 31.9

Qy 417 CGATCTGATATGATTATCATCATGAGACGGCCAGGCTGATGTAGCCGACGCCAGGATTT 47.6

Db 320 TGAATTTTCCACTCATCTGTATTAACAAGGAGGGTTGAAGAAGTGAAGCTGCCGGCT 37.9

Qy 477 CTTCGAGAGATCGTGTCTGCCCTGCACCTACTGCAACGAAACGGGATCGTTCAACGAGA 53.6

Db 380 CTTCAGCAGATTTGTCTGCCGTGACTACTGTCAACAGCAACATGTTGTCCAGAGGA 43.9

Qy 537 TCTCAAGCTGAAAAATCTCTTATAGATGCCAATGAAACATCAAGATTGCTGACTTTGG 59.6

Db 440 CCTGAAGCCAGAAACGTTGCTGGAACGCCACAGATGAATGCTAAGATAGCTGACTTCGG 49.9

Qy 597 CCTCTCCAACTGTACACAAAGGCAATTTCTCCAGAGCTTCTGTGGAGCCCTCTTA 65.6

Db 500 ACTCTCTAATATGATGTCAAGATGGTAATTTTACGAATCTGATGATCCGCAATTA 55.9

Qy 657 CGCCTCGGCTGATAGTCAACAGGGAAGCCATATGCGGCCGAGAGGTAGACAGCTGATC 71.6

Db 560 TGCAGACCCGAGGTCATCTCAGGAAGGCTGATGCGGGTCTCGAGGTGATATCTGGAG 61.9

Qy 717 TCTGGGCGCTTCCTGTACATCTGATGATAGGACCAATGCCCTTTGACGGGCAAGATCA 77.6

Db 620 CTGTGTGTATCTCTGTATGCCCTTCTCTGTGGCACTCCCTTGCAAGATGACACGT 67.9

Qy 777 TAAAACTGTGAAACAATCAGTACGGGGCTTACCGGTGAGCCGCCAAGCCGTCGA 83.6

Db 680 GCGTACGCTCTTTAAGAAGATCGAGGGGGTGTGTTCAATCCCGAATATCTCAACG 73.9

Qy 837 TGCCGTGGC--CTGATCCGCTGCTTTAATGTTGAACCCCAACCCCGTGGGCCACT 89.3

Db 740 TTCTATTCCTCACTCTGTGATGCAATGCTGACGGTGGACCCCTTGAAACGAGCACTAT 79.9

Qy 894 GAGGAGTATGACCACTCAATGGTGGGTCAACTGGGGTTACACA 93.7

Db 800 CAAAGACATACGAGCAGATGAATGTTTAAACAGATTTTGCCA 84.3

```

RESULT 15
US-08-557-006C-38
/ Sequence 38, Application US/08557006C
/ Patent No. 6258547
/ GENERAL INFORMATION:
/ APPLICANT: Berl, Rajindar K.
/ APPLICANT: Carling, David A.
/ TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
/ FILE REFERENCE: NCAP/PM37588/UST
/ CURRENT APPLICATION NUMBER: US/08/557,006C
/ PRIORITY FILING DATE: 1996-03-06
/ PRIOR APPLICATION NUMBER: PCT/GB94/01093
/ PRIOR FILING DATE: 1994-05-20
/ PRIOR APPLICATION NUMBER: GB 9310489.1
/ PRIOR FILING DATE: 1993-05-21
/ PRIOR APPLICATION NUMBER: GB 9318010.7
/ PRIOR FILING DATE: 1993-08-31
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 38
/ LENGTH: 1742
/ TYPE: DNA
/ ORGANISM: Human AMP protein kinase
US-08-557-006C-38

```

Query Match	9.6%	Score 181.2	DB 3	Length 1742
Buett Local Similarity	53.1%	Pred. No. 3.8e-41		
Matches 434	Conservative 0	Mismatches 378	Indels 6	Gaps 2
Oy	126	GGCGGTGAAGGGGACCATCAAAACMACCTCGGCGCATACGATTCCTGGAAC	185	
Db	3	GGCTGAGAACGAGAACCGGCGCGCTGTGATCGGACATCACTACGCTGGGGGAC	62	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 14, 2005, 22:57:45 ; Search time 985.386 Seconds
(without alignments)
11372.259 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 acgagagtcggtgagcttact.....gaatctgcacaaagctcagc 1893

Scoring table: IDENTITY NTC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	4 AAC90433	AAC90433 Murine ly
2	1893	100.0	3073	13 ABD33083	ABD33083 Murine ca
3	1683.4	88.9	2929	6 AAD31710	Aad31710 Rat SNF1
4	1360.2	71.9	1884	12 AD141462	AD141462 Novel hum
5	1360.2	71.9	3353	6 ABK14000	ABK14000 cDNA enco
6	1360.2	71.9	3353	12 AD141460	AD141460 Novel hum
7	1357.6	71.7	3360	6 AAD26459	Aad26459 Human kin
8	1357.6	71.7	3395	4 AAK94280	Aak94280 Human ful
9	1357	71.7	3395	12 ADJ30885	Adj30885 Full leng
10	1357	71.7	3404	13 ABD33085	ABD33085 Human can
11	1357	71.7	3463	12 ADJ96554	Adj96554 Human cal
12	1355.4	71.6	2043	6 ABZ11333	Abz11333 Human pol
13	1355.4	71.6	2501	12 ADM43851	Adm43851 Novel hum
14	1355.4	71.6	3443	5 ABX71420	Abx71420 Human cel
15	1355.4	71.6	3443	10 ADF76964	Adf76964 Novel hum
16	1355.4	71.6	3443	10 ADP81952	Adp81952 Leukemia
17	1355.4	71.6	3443	12 ADO20171	Ado20171 Human pro
18	1338.8	70.7	2291	4 AAF44659	Aaf44659 Novel pro
19	1338.8	70.7	2291	12 AD129357	Ad129357 Human MAR
20	1314.2	69.4	3200	4 AAF75338	Aaf75338 Human TGF

21	1230	65.0	1893	10 AB277163	Ab277163 Human pro
22	1059.4	56.0	37278	13 ABD33082	ABD33082 Murine ca
23	840.2	44.4	2616	6 ABQ72599	Abq72599 Human MDD
24	840.2	44.4	2619	6 ABQ72698	Abq72698 Human MDD
25	723	38.2	39699	13 ABD33084	ABD33084 Human can
26	506.4	26.8	2884	4 AAD14328	Aad14328 Human pro
27	506.4	26.8	6828	4 AAD14327	Aad14327 Human pro
28	506.4	26.8	6828	10 ADE38420	Ad38420 Human pro
29	506.4	26.8	6828	11 ADN95767	Adn95767 Human BGC
30	506.4	26.8	6828	12 ADJ74808	Adj74808 Marker ge
31	506.4	26.8	6828	12 ADL25352	Adl25352 Human ARK
32	506.4	26.8	6828	12 ADQ19733	Adq19733 Human sof
33	506.4	26.8	6828	12 ADP43253	Adp43253 Human pro
34	506.4	26.8	6828	12 ADR25680	Adr25680 Breast ca
35	506.4	26.8	6854	12 ADQ23883	Adq23883 Human sof
36	356.8	18.8	3594	5 AAD03994	Aad03994 Human pro
37	348.8	18.4	587	4 AAK93296	Aak93296 Human CDN
38	348.8	18.4	587	4 AAK91887	Aak91887 Human CDN
39	348.8	18.4	587	12 ADL29723	Adl29723 5' end of
40	348.8	18.4	587	12 ADL28314	Adl28314 5' end of
41	296.8	15.7	1454	4 AB121269	Ab121269 Drosophila
42	267	14.1	1723	4 AAH13802	Aah13802 Human CDN
43	236.6	12.5	1594	4 AAF44655	Aaf44655 Novel pro
44	236.6	12.5	1594	8 AAL60326	Aal60326 Mouse 207
45	236.6	12.5	1594	12 AD129353	Ad129353 Mouse MAR

ALIGNMENTS

RESULT 1	AAC90433	standard; cDNA; 2902 BP.
ID	AAC90433	
XX	AAC90433:	
XX	19-MAR-2001 (first entry)	
XX	Murine Lymph node Stromal cell kinase 1 coding sequence.	
XX		
XX	Murine; Lymph node Stromal cell kinase; MlSK-1; autoimmune disorder;	
XX	wound healing; periodontal disease; inflammatory disease; tumour;	
XX	infection; allergy; sg.	
XX		
OS	Mus musculus.	
XX		
XX	WO200073468-A1.	
XX		
PD	07-DEC-2000.	
XX		
PF	26-MAY-2000; 2000MO-US014696.	
XX		
PR	28-MAY-1999; 99US-0136781P.	
XX		
XX	(IMMV) IMMUNEX CORP.	
XX	Bird TA, Varca GD, Martin U, Anderson DM;	
XX	WPI, 2001-061546/07.	
XX	P-PSDB; AAB50056.	
XX		
PT	Novel murine and human kinase nucleic acids useful for treating	
XX	inflammations, infections, tumors, allergies, autoimmune diseases, and	
XX	for stimulating or suppressing immune responses.	
XX	Claim 1; Page 86-87; 106pp; English.	
XX		
XX	The present sequence is the coding sequence for Murine Lymph node Stromal	
XX	cell kinase 1 (MLSK-1). The protein encoded by the present sequence is	
XX	useful for treating a variety of disorders listed in the disclosure of	
XX	the specification, including autoimmune disorders, allergic reactions,	
XX	myeloid or lymphoid cell deficiencies, wound healing and tissue repair	
XX	and replacement, burns, incisions and ulcers, periodontal disease,	

CC Inflammatory diseases, tumours and bacterial, viral or fungal infection
XX Sequence 2902 BP; 654 A; 794 C; 796 G; 658 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 1893; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGTGGCTTACTTCACAGCGCCGACAGGCTCCTCGGCTTCGCGCTGACC 60
DB 123 ATGAGTCTGGTGGCTTACTTCACAGCGCCGACAGGCTCCTCGGCTTCGCGCTGACC 182

QY 61 TCGGAGAGCGCCGCGCTGCGGACGGGCTCATCAAGTGGCTAAACCTTGATGAAG 120
DB 183 TCGGAGAGCGCCGCGCTGCGGACGGGCTCATCAAGTGGCTAAACCTTGATGAAG 242

QY 121 AAGCAGCGGTGAAGCGGACATCAACAAACAGACCTGCGGACCGCTACGATTCCTG 180
DB 243 AAGCAGCGGTGAAGCGGACATCAACAAACAGACCTGCGGACCGCTACGATTCCTG 302

QY 181 GAGACCTGGGCAAGGCGACCTACGGGAAGTGAAGAAGGACAGAGAGCTCGGGCGT 240
DB 303 GAGACCTGGGCAAGGCGACCTACGGGAAGTGAAGAAGGACAGAGAGCTCGGGCGT 362

QY 241 CTGGTGGCCATCAAGTCCATCAGAAAACAAATATCAAGATGAGAGATCTGTGAC 360
DB 363 CTGGTGGCCATCAAGTCCATCAGAAAACAAATATCAAGATGAGAGATCTGTGAC 422

QY 301 ATACGAGGAGATGATGATCATGTCTTCACTCAACCAACCCACATATGCTCATCAT 360
DB 423 ATACGAGGAGATGATGATCATGTCTTCACTCAACCAACCCACATATGCTCATCAT 482

QY 361 GAAGTGTGAGAAATGACAGCAAGATTGTGATGATGATGATGATGATGATGATGAT 420
DB 483 GAAGTGTGAGAAATGACAGCAAGATTGTGATGATGATGATGATGATGATGATGAT 542

QY 421 CTGTATGATTAATCATCATGAGCGGCAACGGCTGAGAGCGGAGACGACAGGCAATTTCTC 480
DB 543 CTGTATGATTAATCATCATGAGCGGCAACGGCTGAGAGCGGAGACGACAGGCAATTTCTC 602

QY 481 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGACGGATCGTTCAACGAGATCTC 540
DB 603 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGACGGATCGTTCAACGAGATCTC 662

QY 541 AAGCTGAAAAACATCTCTTCTAGATGCCAATGAAACATCAAGTTGCTGACTTTGGCCTC 600
DB 663 AAGCTGAAAAACATCTCTTCTAGATGCCAATGAAACATCAAGTTGCTGACTTTGGCCTC 722

QY 601 TCCAACTGTACCAAAAGGCAAGTCTCTCAGACGTTCTGTGGGAGCCCTCTCTACGCGC 660
DB 723 TCCAACTGTACCAAAAGGCAAGTCTCTCAGACGTTCTGTGGGAGCCCTCTCTACGCGC 782

QY 661 TCGCCTGAGATGATCAACGGGAAGCCCTATGTTGGGCGCAGAGGTGACAGCTGCTCTG 720
DB 783 TCGCCTGAGATGATCAACGGGAAGCCCTATGTTGGGCGCAGAGGTGACAGCTGCTCTG 842

QY 721 GGGCTTCTCTGTACATCTGTGTGATGAGCAATGCGCTTTGACGCGGACAGATCAATA 780
DB 843 GGGCTTCTCTGTACATCTGTGTGATGAGCAATGCGCTTTGACGCGGACAGATCAATA 902

QY 781 ACACTGTGAAGCAATCAAGTACGGGGCTTACCGTGAAGCGGCGCAAGCGCTCCGATGCGC 840
DB 903 ACACTGTGAAGCAATCAAGTACGGGGCTTACCGTGAAGCGGCGCAAGCGCTCCGATGCGC 962

QY 841 TGTGGCTGATCCGATGCTGTATGATGTAACCCACCGCTGAGGCGACACTGAGAGAT 900
DB 963 TGTGGCTGATCCGATGCTGTATGATGTAACCCACCGCTGAGGCGACACTGAGAGAT 1022

QY 901 GTAGCCAGTCAATGATGAGTCACTGGGGTTACACCAACCGGATCGGAGAACAGAAAGCC 960
DB 1023 GTAGCCAGTCAATGATGAGTCACTGGGGTTACACCAACCGGATCGGAGAACAGAAAGCC 1082

QY 961 CTGCGTGAAGGTGGGCAACCTAGTGTGACTTTGGCGGGCTTCATGGCGGACGTGTTA 1020

DB 1083 |||||CTGCTGAAGGTGGGCAACCTTAGTGTGACTTTGGCGGGCTTCATGGCGGACTGTTA 1142

QY 1021 CGTGCCTCTCGCGCCCTCTCTGAGAAATGAGCCAGAGTGTGACGTTCTTCAAGCAG 1080

DB 1143 CGTGCCTCTCGCGCCCTCTCTGAGAAATGAGCCAGAGTGTGACGTTCTTCAAGCAG 1202

QY 1081 CAGTGCCTGGAGGTGGAAGCACTGTACTGCGGCTGAGACGGGCAATTTCTTAAAG 1140

DB 1203 CAGTGCCTGGAGGTGGAAGCACTGTACTGCGGCTGAGACGGGCAATTTCTTAAAG 1262

QY 1141 TCCGAAAGAGATGATCATGCTGCTGAAATCTGCAAGGATCCCGCTGAGATACCTCT 1200

DB 1263 TCCGAAAGAGATGATCATGCTGCTGAAATCTGCAAGGATCCCGCTGAGATACCTCT 1322

QY 1201 TCTCGCCTGGCAAGACGACCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAAGTCTCT 1260

DB 1323 TCTCGCCTGGCAAGACGACCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAAGTCTCT 1382

QY 1261 ACTTGTCAAGGAGGTACAGAGACCTTCAAGAACTCAAGACCGGTGCTGATCTCA 1320

DB 1383 ACTTGTCAAGGAGGTACAGAGACCTTCAAGAACTCAAGACCGGTGCTGATCTCA 1442

QY 1321 GGGCAGCCTGTCCCTGTGATCCCTGCTCCAGAGGAAAGGATCCTTAAAGTCTGCA 1380

DB 1443 GGGCAGCCTGTCCCTGTGATCCCTGCTCCAGAGGAAAGGATCCTTAAAGTCTGCA 1502

QY 1381 CAGCGTGAATCTGTACTACTCTCTCCAGAGCCAGGAGTCTGGGAACTTTAGAC 1440

DB 1503 CAGCGTGAATCTGTACTACTCTCTCCAGAGCCAGGAGTCTGGGAACTTTAGAC 1562

QY 1441 GCCAGTATGTGTTGTGATGAGGAGCCCGGTGAGACAGAGTCTTCAACAGCTTCAAGG 1500

DB 1563 GCCAGTATGTGTTGTGATGAGGAGCCCGGTGAGACAGAGTCTTCAACAGCTTCAAGG 1622

QY 1501 CTCTCTCTCAACCGGCAAGGCAATTTCTCAAACTCAATGGAAGTTCTCCCGCAAGCTTA 1560

DB 1623 CTCTCTCTCAACCGGCAAGGCAATTTCTCAAACTCAATGGAAGTTCTCCCGCAAGCTTA 1682

QY 1561 GAAGCAATACCCCTAGACACTTTGGCTCTCCAGACCAACTGAGCTCTCCATCCTGCA 1620

DB 1683 GAAGCAATACCCCTAGACACTTTGGCTCTCCAGACCAACTGAGCTCTCCATCCTGCA 1742

QY 1621 GCCCGGCCAGCCGCCCTCAGGGGCTGTGATGAGAGACAGCATCTGCTCCAGATCC 1680

DB 1743 GCCCGGCCAGCCGCCCTCAGGGGCTGTGATGAGAGACAGCATCTGCTCCAGATCC 1802

QY 1681 TTTGACCAATGGAATGTGCTGAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1740

DB 1803 TTTGACCAATGGAATGTGCTGAAGTCTTCCGAAACCCCACTGAGGGGCTGTGTCT 1862

QY 1741 GTGACAACTGAGGGGCTTGAAGCAGCTCCCTGAGAAAGTCTGAAGCGATGTGGCAG 1800

DB 1863 GTGACAACTGAGGGGCTTGAAGCAGCTCCCTGAGAAAGTCTGAAGCGATGTGGCAG 1922

QY 1801 GAATCTTGGGAGATGCTGCTTTCTGTGACAGCTGCAAGAGGTGACTGCAAGCTTAC 1860

DB 1923 GAATCTTGGGAGATGCTGCTTTCTGTGACAGCTGCAAGAGGTGACTGCAAGCTTAC 1982

QY 1861 AGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893

DB 1983 AGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 2015

RESULT 2
ABD33083
ID ABD33083 standard; cDNA; 3073 BP.
XX
XX ABD33083;
AC
XX
DT 18-NOV-2004 (first entry)
XX
DE Murine cancer-associated (CA) cDNA M807-004.

XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW ss; cancer; cystostatic.
XX Mus musculus.
XX WO2004058146-A2.
XX 15-JUL-2004.
XX 15-DEC-2003; 2003WO-US040081.
XX 17-DEC-2002; 2002US-00322281.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX MPI; 2004-499109/47.
XX P-PSDB; AB084756.
XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX Disclosure; SEQ ID NO 22; 182bp; English.
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
XX associated (CA) nucleic acids encoding them. The invention also relates
XX to a method for treating cancers involving administering to a patient an
XX inhibitor of CAP, and a method of screening for anticancer activity in a
XX potential drug involving providing a cell that expresses a CA gene,
XX contacting a tissue sample derived from a cancer cell with an anticancer
XX drug candidate and monitoring the effect of the anticancer drug candidate
XX on expression of the CA gene. The CAP proteins are useful for detecting
XX cancer associated with expression of a CAP protein in a test cell sample
XX and for screening for a bioactive agent capable of modulating the
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX cancer, involving determining the expression of a CA nucleic acid in a
XX tissue. This sequence represents murine CA cDNA of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 3073 BP; 690 A; 827 C; 844 G; 712 T; 0 U; 0 Other;
Query Match 100.0%; Score 1893; DB 13; Length 3073;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAATCGGTGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 60
DB 92 ATGGAATCGGTGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 151
QY 61 TCGGAGAGCGCCGCGCTGCGGACGCGCTCATCAAGTCGCTAAACCTTGATGAAG 120
DB 152 TCGGAGAGCGCCGCGCTGCGGACGCGCTCATCAAGTCGCTAAACCTTGATGAAG 211
QY 121 AAGCAGCGGTGAGAGCGGACCATCAAAACACACCTGCGGACCGCTACGAGTTCCTG 180
DB 212 AAGCAGCGGTGAGAGCGGACCATCAAAACACACCTGCGGACCGCTACGAGTTCCTG 271
QY 181 GAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGACAGAGAGACTGGGGCGCT 240
DB 272 GAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGACAGAGAGACTGGGGCGCT 331
QY 241 CTGGTGGCCATCAAGTCATCAGAAAGCAAAATCAAGATGAGAGATTCGCTGAC 300
DB 332 CTGGTGGCCATCAAGTCATCAGAAAGCAAAATCAAGATGAGAGATTCGCTGAC 391
QY 301 ATACGAGGAGATGATGATCATGTCTTCACTCAACCAACCCCAATTCATTCATCAT 360
DB 392 ATACGAGGAGATGATGATCATGTCTTCACTCAACCAACCCCAATTCATTCATCAT 451

QY 361 GAAAGTTTGAGAAATGACGAAAGATTGATTTGTATGATGAGATATGCAAGCCGAGT 420
DB 452 GAAAGTTTGAGAAATGACGAAAGATTGATTTGTATGATGAGATATGCAAGCCGAGT 511
QY 421 CTGTATGATTTAATCAATGATGAGGCGGACGCTGAGTGAAGGAGGACGAGCAATTCCT 480
DB 512 CTGTATGATTTAATCAATGATGAGGCGGACGCTGAGTGAAGGAGGACGAGCAATTCCT 571
QY 481 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGAGCGGATGCTTCAACGAGATCTC 540
DB 572 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGAGCGGATGCTTCAACGAGATCTC 631
QY 541 AAGCTGAAAACATCTCTTCTAGATGCAATGAAATCAATCAAGATTGCTGACTTGGCTC 600
DB 632 AAGCTGAAAACATCTCTTCTAGATGCAATGAAATCAATCAAGATTGCTGACTTGGCTC 691
QY 601 TCCAACTGTATCAACAAAGGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTACGCGC 660
DB 692 TCCAACTGTATCAACAAAGGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTACGCGC 751
QY 661 TCGCCTGAGTATGATCAACGGAAGCCCTATGTTGGGCGCAGAGGTGACAGCTGCTCTG 720
DB 752 TCGCCTGAGTATGATCAACGGAAGCCCTATGTTGGGCGCAGAGGTGACAGCTGCTCTG 811
QY 721 GAGCTTCTCTGTATCACTCTGTATGATGCAACCATGCTCTTTGACGCGGACAGATCAATAA 780
DB 812 GAGCTTCTCTGTATCACTCTGTATGATGCAACCATGCTCTTTGACGCGGACAGATCAATAA 871
QY 781 ACACTGATGAGCAAAATCAATGATGAGGCGCTTACCTGAGCGCCCAAGCCTGCGATGCC 840
DB 872 ACACTGATGAGCAAAATCAATGATGAGGCGCTTACCTGAGCGCCCAAGCCTGCGATGCC 931
QY 841 TGTGGGCTATTCGGGTGCTTTAATGATGAAACCCCAACCGGTGGGCAACATGAGGAT 900
DB 932 TGTGGGCTATTCGGGTGCTTTAATGATGAAACCCCAACCGGTGGGCAACATGAGGAT 991
QY 901 GTAGCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 992 GTAGCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
QY 961 CTGCTGAGGAGTGGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1052 CTGCTGAGGAGTGGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
QY 1021 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGAGGCTTCTTCAAGCAG 1080
DB 1112 CGTGGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGAGGCTTCTTCAAGCAG 1171
QY 1081 CAGTGCAGGAGTGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1172 CAGTGCAGGAGTGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
QY 1141 TCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1232 TCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
QY 1201 TCTGCGCTGCGCAAGGACGCTTAAAGCTTCCGAAAGGCACTTCTCAAGAAAAAAGTCTCT 1260
DB 1292 TCTGCGCTGCGCAAGGACGCTTAAAGCTTCCGAAAGGCACTTCTCAAGAAAAAAGTCTCT 1351
QY 1261 ACCTGCTCAAGGAGGATGAGAGGACCTTCAAGAACTCAAGACCGGCTGATATCTCA 1320
DB 1352 ACCTGCTCAAGGAGGATGAGAGGACCTTCAAGAACTCAAGACCGGCTGATATCTCA 1411
QY 1321 GGGCAGCCTGTCTCTGCTGATCTCTGCTCCCAAGAAAGGCACTCTTAAGAAATCTCGA 1380
DB 1412 GGGCAGCCTGTCTCTGCTGATCTCTGCTCCCAAGAAAGGCACTCTTAAGAAATCTCGA 1471
QY 1381 CAGCGTGAATCTGATTAATCTCTCTCTCAAGACCCAGAGCTGTGGGAACTCTTAAGAC 1440
DB 1472 CAGCGTGAATCTGATTAATCTCTCTCTCTCAAGACCCAGAGCTGTGGGAACTCTTAAGAC 1531
QY 1441 GCCAGTGAATGTGTTTGAAGTGGGACCCCGTGAAGAGAAATCTCCACAGGCTTCAGGG 1500

DB 1532 GCCAGGATGATGTTTGTGATGGGGACCCGCTGGAGAGAACTCTCCACAGGCTTCAAGG 1591
QY 1501 CTCTCTCTCCACCGCAAGGGCAATTTCTCAAACTCAATGGCAAGTTTCCCGCAAGCCTTA 1560
DB 1592 CTCTCTCTCCACCGCAAGGGCAATTTCTCAAACTCAATGGCAAGTTTCCCGCAAGCCTTA 1651
QY 1561 GAAGGACATACCCCTGACATCTTTGAGTCCCTGGACCAATGGCTCTCCCATCTCCGA 1620
DB 1652 GAAGGACATACCCCTGACATCTTTGAGTCCCTGGACCAATGGCTCTCCCATCTCCGA 1711
QY 1621 GCCCGGCCGACCGCCCTCTGAGGGGCTGTGAGTGAAGACAGATCTCTCTCCGAGTCC 1680
DB 1712 GCCCGGCCGACCGCCCTCTGAGGGGCTGTGAGTGAAGACAGATCTCTCTCCGAGTCC 1771
QY 1681 TTTGACCAATTGACATTTGCTGCTGAAAGCTTTCCGAAAACCCCACTGAGGGGCTGTGCT 1740
DB 1772 TTTGACCAATTGACATTTGCTGCTGAAAGCTTTCCGAAAACCCCACTGAGGGGCTGTGCT 1831
QY 1741 GTGGAACAACCTGAGGGGGCTTGAAGAGCCTCCCTGAGAAAGTCTGAAAGCATGTGGCAG 1800
DB 1832 GTGGAACAACCTGAGGGGGCTTGAAGAGCCTCCCTGAGAAAGTCTGAAAGCATGTGGCAG 1891
QY 1801 GAATCCTTGGGGGATGAGTCTCTTTCTCTGACAGACTGCGAAGAGTGACTGCAAGCTTAC 1860
DB 1892 GAATCCTTGGGGGATGAGTCTCTTTCTCTGACAGACTGCGAAGAGTGACTGCAAGCTTAC 1951
QY 1861 AGACAAAGCCTTAGGAATCTGCTCAAAAGCTCAGC 1893
DB 1952 AGACAAAGCCTTAGGAATCTGCTCAAAAGCTCAGC 1984
RESULT 3
AAD31710
ID AAD31710 standard; cDNA; 2929 BP.
XX AAD31710;
AC
XX
DT 18-JUN-2002 (first entry)
DE Rat SNF1/AMPK-Related Kinase (SNARK) cDNA.
XX
XX Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;
KM glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;
KM hyperglycaemic; drug screening; hypoglycaemia; ss.
XX
OS Rattus sp.
XX
XX
FH Key Location/Qualifiers
FT CDS 83..1975
FT /tag= a
FT /product= "Rat SNARK protein"
PN WO200212456-A2.
XX
PD 14-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-CA001109.
XX
XX 03-AUG-2000; 2000US-022265OP.
PR 12-MAR-2001; 2001US-0274613P.
PR 28-MAR-2001; 2001CA-02340783.
XX
XX (ONMO-) 1149336 ONTARIO INC.
XX
XX PA
XX PI Drucker DJ, Rosen CF, Lefebvre DL,
XX
XX MPI: 2002-241747/29.
XX
XX P-PSDB; AAE19885.
XX
XX AMPK (AMP-activated protein kinase)-related kinase, designated SNARK
PT polypeptides and polynucleotides, useful for treating or preventing
PT diabetes, or other disorders of lipoprotein production leading to

PT Increased levels of cholesterol.
XX
XX Example; Fig 2; 94pp; English.
XX
XX The invention relates to an AMPK (AMP-activated protein kinase)-related
CC kinase, designated SNARK polypeptides and polynucleotides. SNARK
CC (SNF1/AMP-activated protein kinase) is involved in stress response to
CC glucose deprivation. The polynucleotides are useful for expressing SNARK
CC protein in isolated form or as a protein conjugate. Activation of SNARK
CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
CC in other cell types such as heart and skeletal muscles, as well as
CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
CC to have insulin-like effects that would enhance the disposal of glucose
CC into muscle and reduce plasma glucose for the treatment of diabetes and
CC some type of disorders of lipoprotein production leading to increased
CC levels of cholesterol or triglycerides. SNARK or its variants may be
CC administered to a subject to treat or prevent a disease associated with
CC decreased expression of SNARK, such as diabetes. SNARK antibodies are
CC used to modulate SNARK activity either in vivo for therapeutic purposes,
CC or in vitro, for drug screening and related investigational purposes.
CC SNARK antagonists may be administered to increase fuel production,
CC decrease glucose uptake and increase levels of blood glucose in a patient
CC suffering from hypoglycaemia. The present sequence is rat SNARK cDNA
XX
SQ Sequence 2929 BP; 650 A; 788 C; 791 G; 700 T; 0 U; 0 Other;
Query Match 88.9%; Score 1683.4; DB 6; Length 2929;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
QY 1 ATCGAATCGGTGGCTTACTCCAGCGCCGAGCGAGGCTTCCCTCGGCTCCGCTGGCC 60
DB 83 ATGGAATCGGTGGCTTACTCCAGCGCCGAGCGAGGCTTCCCTCGGCTCCGCTGGCC 142
QY 61 TCAGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTCGCTAACTCGATGAAG 120
DB 143 ACGGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTCGCTAACTCGATGAAG 202
QY 121 AAGCAGGCGGTGAAGCGGACCATCAAAACAACCTCGGACCGCTACGAGTTCCTG 180
DB 203 AAGCAGGCGGTGAAGCGGACCATCAAAACAACCTCGGACCGCTACGAGTTCCTG 262
QY 181 GAGACGCTGGGCAAGGGCACTTACGGGAAAGTGAAGAGCAGAGAGCTTGGGGCT 240
DB 263 GAGACGCTGGGCAAGGGCACTTACGGGAAAGTGAAGAGAGCAGAGAGCTTGGGGCA 322
QY 241 CTGGTGGCATCAAGTCATCAAGGAAAGCAAAATCAAGATGAGAGATCTGCGGAC 300
DB 323 CTGGTGGCATCAAGTCATCAAGGAAAGCAAAATCAAGATGAGAGATCTGCGGAC 382
QY 301 ATACGAGGAGATGAGATCAATGCTTCACTCAACCAACCCCATCATTTGCCATCAT 360
DB 383 ATACGAGGAGATGAGATCAATGCTTCACTCAACCAACCCCATCATTTGCCATCAT 442
QY 361 GAAGTCTTGAATGACAGCAAGATTTGATTTGATGATGATGATGATGATGATGATGAT 420
DB 443 GAAGTCTTGAATGACAGCAAGATTTGATTTGATGATGATGATGATGATGATGATGAT 502
QY 421 CTGTATGATTAATCAATGAGGAGGAGCGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGG 480
DB 503 CTGTATGATTAATCAATGAGGAGGAGCGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGG 562
QY 481 CGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 563 CGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
QY 541 AAGCTGGAAGAAATCTTTAGATGCAATGGAACATCAATGATGATGATGATGATGATGAT 600
DB 623 AAGCTGGAAGAAATCTTTAGATGCAATGGAACATCAATGATGATGATGATGATGATGAT 682
QY 601 TCCAACTGTATCAAAAGGCAAGTTCCTCCAGAGAGTCTGTGGAGAGCCCTCTACAGCC 660
DB 683 TCCAACTGTATCAAAAGGCAAGTTCCTCCAGAGAGTCTGTGGAGAGCCCTCTATAGCC 742

QY 661 TCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGAACAGCTGCTCTG 720
DB 743 TACCTGAGATGCTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGAACAGCTGCTCTG 802
QY 721 GGGGTTCTCTGTGACATCTGTGTGACATGAGCAACATGCTTTTGAACGGGCAAGATCATAA 780
DB 803 GGGGTTCTCTGTGACATCTGTGTGACATGAGCAACATGCTTTTGAACGGGCAAGATCATAA 862
QY 781 ACACGTGTGAAGAAATCAGTAACGGGGCTTACCGTGTAGCGGCCCAAGCGCTCGATGCC 840
DB 863 ACCCTGTGAACAAATCAGTAAGGGGCTTACCGAGAGCGGTGCAAAACGCTGTGATGCC 922
QY 841 TGTGGCTGTATCCGCTGTGTATGTGTAAATGTGAAACCCCAACCGGTGGGCAACATGAGAT 900
DB 923 TGTGGCTGTATCCGCTGTGTATGTGTAAATGTGAAATCCCAACCGGTGGGCAACATGAGAT 982
QY 901 GTAGCCATGATGTTGTGTGTCAACTGTGGGTTTACACCAACCGAGTGGGGAAACAGAAAGCC 960
DB 983 GTAGCCATGATGTTGTGTGTCAACTGTGGGTTTACACCAACCGAGTGGGGAAACAGAAAGCT 1042
QY 961 CTGCGTGAAGGGTGGGACCCCTAGTGTGACTTTGGCGGGGCTCCAGTGGGGGCTGCTGTA 1020
DB 1043 CTGCGTGAAGGGTGGGACCCCTAGTGTGACTTTGGCGGGGCTCCATGTGGGCTGCTGTA 1102
QY 1021 CGTGGCTCTCTGCGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTTCAAGCAG 1080
DB 1103 CGTGGCTCTCTGCGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTTCAAGCAG 1162
QY 1081 CAGGTGCGGGAGTGTGAGCACTGTATCTGTGGCTGTGAGGGGCAACATTTCTTAAAGAG 1140
DB 1163 CATGTGCGGGAGTGTGAGCACTGTATCTGTGGCTGTGAGGGGCAACATTTCTTAAAGAG 1222
QY 1141 TCCCGAAAGAAATGATGCTCAAAATCTGCAAGGTGACCCGCTGTGAGTACTCT 1200
DB 1223 TCCCGAAAGAAATGATGCTCAAAATCTGCAAGGTGACCCGCTGTGAGTACTCT 1282
QY 1201 TCTGCGCTGTGCAAGAGCAAGCTTTAAGCTTCCGAAAGGCAATTTCAAGAAAGCTCT 1260
DB 1283 TCTGCGCTGTGCAAGAGCAAGCTTCAAGCTTCCGAAAGGCAATTTCAAGAAAGGCTCT 1342
QY 1261 ACCTGTGTGAGGGAGGTGACAGAGAGCCCTCAGAGAACTCAGACCGGCTGTGATCTCA 1320
DB 1343 CCTCATGCGGGAGGTGACAGAGAGCCCTCAGAGAACTCAGAGCGGTGTCAATATCCCA 1402
QY 1321 GGGGACCTGTCTCTGTGTATCTCTGTGCAAGGCCAGAGGCAATCTTAAAGTCTCGA 1380
DB 1403 GGGGACCTGTCTCTGTGTATCTCTGTGCAAGGCCAGAGGCAATCTTAAAGTCTCGG 1462
QY 1381 CAGCGTGAATCTGTGTATCTCTGTGCAAGGCCAGAGGCAATCTTAAAGTCTCGA 1440
DB 1463 CAGCGTGAATCTGTGTATCTCTGTGCAAGGCCAGAGGCAATCTTAAAGTCTCGA 1522
QY 1441 GCCAGT 1500
DB 1523 GCAAGT 1582
QY 1501 CTCTCTCTGT 1560
DB 1583 ---CGCTCTCATGTGCAAGGGCAATCTCAAACTCAATGTGCAAGTGTGTGTGTGTGT 1639
QY 1561 GAAGGCACTACCCCTGTGCACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
DB 1640 GAAGGCACTACCCCTGTGCACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1699
QY 1621 GCCCGGCGCAGCGCCCTCAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
DB 1700 GCCCGGCGCAGCGCCCTCAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1759
QY 1681 TTTGACCAATTTGACTTGT 1740
DB 1760 TTTGACCAATTTGACTTGT 1819

QY 1741 GTGAGCACTGTGAGGGGCTTTGAGCAAGCTCTCTCAGAAAGTGTGAAGCGATGTGGCAG 1800
DB 1820 GTGAGCACTGTGAGGGGCTTTGAGCAAGCTCTCTCAGAAAGTGTGAAGCGATGTGGCAG 1879
QY 1801 GAATCTGTGGGGATATGCTGCTTTCTGTGACAGACTGCCAAGAGGTGACTGGAGCCTAC 1860
DB 1880 GAATCTGTGGGGATATGCTGCTTTCTGTGACAGACTGCCAAGAGGTGACTGGAGCCTAC 1939
QY 1861 AGACAAGCCCTAGGAATCTGCTCAAGCTGAGC 1893
DB 1940 AGACAAGCCCTAGGAATCTGCTGAGACTGAGC 1972

RESULT 4
ADL14162
ID ADL14162 standard; cdna; 1884 BP.
XX
AC ADL14162;
XX
DT 17-JUN-2004 (first entry)
XX
DE Novel human gene 3700 coding region.
KW cytostatic; cardiant; hypotensive; antianginal; osteopathic;
KW antiarthritic; antirheumatic; neuroprotective; antiinflammatory;
KW antipsoriatic; antiaslomatic; cardiovascular; vitruce; analgesic; CNS;
KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;
KW nephrotropic; antithyroid; dermatological; immunomodulator;
KW cell proliferation disorder; cell differentiation disorder;
KW brain disorder; platelet disorder; breast disorder; colon disorder;
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;
KW thyroid disorder; testes disorder; haematopoietic disorder;
KW pancreatic disorder; skeletal muscle disorder; skin disorder;
KW dermal disorder; bone metabolism disorder; immune disorder;
KW inflammatory disorder; cardiovascular disorder;
KW endometrial cell disorder; liver disorder; viral disease; pain disorder;
KW metabolic disorder; neurological disorder;
KW central nervous system disorder; erythroid disorder;
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
KW hypertension; angina; osteoarthritis; Rheumatoid arthritis;
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
KW cell proliferation; cell differentiation; cell growth; cell division;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
XX
PN US2004058355-A1.
XX
PD 25-MAR-2004.
XX
PF 25-APR-2003; 2003US-00423543.
XX
XX 30-SEP-1998; 98US-00163821.
PR 27-JAN-1999; 99US-0117580P.
PR 25-MAR-1999; 99US-00276400.
PR 30-JUL-1999; 99US-00365162.
PR 09-SEP-1999; 99US-00392189.
PR 05-OCT-1999; 99US-00412210.
PR 23-NOV-1999; 99US-00448076.
PR 29-FEB-2000; 2000US-0186061P.
PR 28-APR-2000; 2000US-0206688P.
PR 19-MAY-2000; 2000US-0205447P.
PR 30-JUN-2000; 2000US-00608921.
PR 31-JUL-2000; 2000US-0221925P.
PR 25-SEP-2000; 2000US-0234922P.
PR 25-SEP-2000; 2000US-0235035P.
PR 08-NOV-2000; 2000US-0246669P.
PR 09-NOV-2000; 2000US-00711216.
PR 14-NOV-2000; 2000US-0248325P.
PR 15-NOV-2000; 2000US-0248893P.
PR 22-DEC-2000; 2000US-0257511P.
PR 05-JAN-2001; 2001US-0260166P.

PR 28-FEB-2001; 2001US-00797039.
PR 27-FEB-2001; 2001US-00845044.
PR 20-JUL-2001; 2001US-00909743.
PR 31-JUL-2001; 2001US-00920346.
PR 13-AUG-2001; 2001US-00928531.
PR 14-AUG-2001; 2001US-00929218.
PR 15-AUG-2001; 2001US-0312539P.
PR 25-SEP-2001; 2001US-00963159.
PR 08-NOV-2001; 2001US-0008016.
PR 13-NOV-2001; 2001US-00012055.
PR 15-NOV-2001; 2001US-00003690.
PR 30-JAN-2002; 2002US-00060763.
PR 25-MAR-2002; 2002US-00105989.
PR 12-APR-2002; 2002US-00121911.
PR 12-APR-2002; 2002US-00217168.
PR 22-OCT-2002; 2002US-00278036.
PR 02-JAN-2003; 2003US-00336489.
PR 03-JAN-2003; 2003US-00336153.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PI Kapeller-Liebermann R, Hunter JI, Meyers RE, Rudolph-Owen LA;
PI Curtis RAJ, Olandt PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;
PI Silos-Santiago I, Bandaru R;
XX
XX MPI: 2004-268788/25.
DR P-PSDB; ADL4161.
XX
XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,
PT 26176, 26343, 56638, 18610, 33217, 21967, h1983, 38555 or 593
PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,
XX heart failure and angina.
XX
XX Claim 1; SEQ ID NO 45; 139pp; English.
XX
XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,
CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising
CC any one of 40 nucleotide sequences (I). The nucleic acid molecules and
CC polypeptides are useful for diagnosing and treating a subject having a
CC disorder, or a subject at risk of developing a disorder, which is
CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,
CC m1983, 38555 or 593 activity, such as cellular proliferation and/or
CC differentiative disorders, brain disorders, platelet disorders, breast
CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
CC ovarian disorders, prostate disorders, cervical disorders, spleen
CC disorders, thymus disorders, thyroid disorders, testis disorders,
CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
CC skin (dermal) disorders, disorders associated with bone metabolism,
CC immune, e.g. inflammatory disorders, cardiovascular disorders,
CC endothelial cell disorders, liver disorders, viral diseases, pain
CC disorders, metabolic disorders, neurological or central nervous system
CC disorders, erythroid disorders, blood vessel disorders or angiogenic
CC disorders (all claimed), e.g. cancer, heart failure, hypertension, Crohn's
CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
CC disease, psoriasis, or asthma. The nucleic acid molecules and
CC polypeptides are also useful as modulating agents in regulating a variety
CC of cellular processes, e.g. cell proliferation, differentiation, growth and
CC division. This sequence encodes a novel human protein of the invention.
CC Note: The sequences given in the specification are also available in
CC electronic format from
CC ftp.segdata.nebplc.gov/sequence.html?DocID=20040058355.
XX
XX Sequence 1884 BP; 403 A; 599 C; 548 G; 334 T; 0 U; 0 Other;
SQ

Query Match 71.9%; Score 1360.2; DB 12; Length 1884;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 155; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

QY 47 CCTCCGCTGCTCGAGAGCGCGCGCTGCGAGCGGCTCATCAATGCGGCA 106
DB 35 CCACTCCCTCGGCGGAGAGCTAGCCCGCGCTGCGAGAGGCGCTGATCAAGTCCGCA 94

QY 107 AACCTGTATGAAGAGAGCGCGGTGAAGCGGACCATCAACAACAACTGCGGACC 166
DB 95 AGCCCTTAATGAAGAGAGAGCGGCGGTGAAGCGGACCATCAACAACAACTGCGGACC 154
QY 167 GCTACAGATTCCTGAGAGCGCGGTGAAGAGCGGCGGCTTACGAGAGGTGAAGAGCGAG 226
DB 155 GCTACAGATTCCTGAGAGCGCGGTGAAGAGCGGCGGCTTACGAGAGGTGAAGAGCGAG 214
QY 227 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 286
DB 215 AGAGCTCGGGGCGCGCTGTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGATGAGC 274
QY 287 AGGATCTGTGCAATATGAGAGAGGAGATTAATGATCATGTTCTTCACTCAACACCCGACA 346
DB 275 AAGATCTGATGCAATATGAGAGAGGAGATTAATGATCATGTTCTTCACTCAACACCCGACA 334
QY 347 TCATTGCCATCCATGAAGTGTGAAGAAATGACGCAAGATTGATGATGATGATGATG 406
DB 335 TCATTGCCATCCATGAAGTGTGAAGAAATGACGCAAGATTGATGATGATGATGATG 394
QY 407 CCAGCGAGGCGCATCTGTATGATTAATCATCATGAGCGGCGGCGTGAAGCGGAGC 466
DB 395 CCAGCGGCGGCGCATCTGTATGATTAATCATCATGAGCGGCGGCGTGAAGCGGAGC 454
QY 467 CCAGGCAATTTCTTCGACAGATGATGATGATGATGATGATGATGATGATGATGATG 526
DB 455 CTAGGCAATTTCTTCGACAGATGATGATGATGATGATGATGATGATGATGATGATG 514
QY 527 TTCACCGAATCTCAAGCTGAGAAATATCTCTTGAATGCAATGAAATCAATGATG 586
DB 515 TCACCGAATCTCAAGCTGAGAAATATCTCTTGAATGCAATGAAATCAATGATG 574
QY 587 CTGACTTGTGCTCTCAACATGATCAACAAAGGCAAGTTCCTCCAGAGCTTGAGGA 646
DB 575 CTGACTTGTGCTCTCAACATGATCAACAAAGGCAAGTTCCTCCAGAGCTTGAGGA 634
QY 647 GCGCTTCTACGCGCTGCGCTGAGATAGTCAAGGAGGCGGCTATGAGGCGGAGAGTGG 706
DB 635 GCGCTTCTACGCGCTGCGCTGAGATAGTCAAGGAGGCGGCTATGAGGCGGAGAGTGG 694
QY 707 ACAAGCTGTCTCTGCGGCTTCTCTGATCATCTGATGATGATGATGATGATGATGATG 766
DB 695 ACAAGCTGTCTCTGCGGCTTCTCTGATCATCTGATGATGATGATGATGATGATGATG 754
QY 767 GCGAGGATCATTAACATCTGAGCAATGATCAAGGAGGCTTCAAGGAGGCGGCGCA 826
DB 755 GCGAGGATCATTAACATCTGAGCAATGATCAAGGAGGCTTCAAGGAGGCGGCGCA 814
QY 827 AGCGCTCGATGCTGTGCGCTGATCGGCTGATCGGCTGATCGGCTGATCGGCTGATCG 886
DB 815 AACCTCTGATGCTGTGCGCTGATCGGCTGATCGGCTGATCGGCTGATCGGCTGATCG 874
QY 887 CCAACTGAGAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 875 CCAACTGAGAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 934
QY 947 GGAACAGAGAGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1006
DB 935 GGAACAGAGAGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 994
QY 1007 TGCGGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 995 TGCGGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1054
QY 1067 GCTTCTTCAAGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1126
DB 1055 GCTTCTTCAAGAGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
QY 1127 ATTCTCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1186
DB 1115 ATTCTCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1174

QY 1187 CTGAGATACCTCTCTCGCCCTGGCAAGACGCTTAAGCTTCGAAAGGCAATCTCA 1246
 |||||
 Db 1175 CTGATGACACTGCGCCCTGGCAAGACCACTCAAGCTGCCAAAGGCAATCTCA 1234
 |||||
 QY 1247 AGAAAGTCTCTACTCTCTGAGGAGGTACAGAGGACCTCGAGAACTGAGACCGG 1306
 |||||
 Db 1235 AGAAGAGGTGTGAGGCTCTGCAAGAGGAGGTACAGAGGACCTCGAGCTCAAGCCAA 1294
 |||||
 QY 1307 TGCCTATACTCCAGGAGCAGCTGTCTCTGTATCTCCCTGCCAAGAAAGGCAATCC 1366
 |||||
 Db 1295 TCCCTCGAGCCAGGAGGAGGCTGCCCC-----CCCTGCTCCCAAGAGGCAATCC 1345
 |||||
 QY 1367 TTAAGAGTCTGACAGCCGTAATCTGTGTTACTACTCTCTCGAAGCCAGGAGTCTG 1426
 |||||
 Db 1346 TCAAGAGCCCGCAGAGCCGAGCTGTGCTACTACTCTCTCCGAGCCAGTGAATCTG 1405
 |||||
 QY 1427 GGGAACTCTTAAAGCCGAGTATGTGTTGTGAGTGGGAGCCCGGTGAGACAGAACTCTC 1486
 |||||
 Db 1406 GGGAGCTCTTGAAGCAGAGGCGAGGTGTGTGTGAGTGGGAGTCCCAAGAGCAGAAAGCTTC 1465
 |||||
 QY 1487 CACAGGCTTCAAGGAGCTCTCTCCACAGGAGGCAATTCCAACTCAATGGGCAAGTTC 1546
 |||||
 Db 1466 CGAAGCTTCAAGGAGCTCTCTCCATGCAAGGCAATCTCAACTCAATGGGCAAGTTC 1525
 |||||
 QY 1547 CCCGCAAGCTTAAAGAGCACTACCCCTAGACACTTTCCTCTGACCAACTGAGCTT 1606
 |||||
 Db 1526 CCCAGACAGCTTAAAGAGCTGCGGAGCCCAACACTTTCCTCTGAGTGAATGAGCTGCC 1585
 |||||
 QY 1607 CTTCCATCTTGAAGCCGAGCCGAGCCCTCTGAGGAGCTGTGAGTGAAGACAGATCC 1666
 |||||
 Db 1586 CACTCGCCCTGAGCCGAGCCGAGCCGAGCTGAGGAGGAGAGACAGATCC 1645
 |||||
 QY 1667 TGTCTCCAGTCTTGAAGCAATGAGCTGTGCTGAGAGCTCTTCCGAAAGCCCACTGA 1726
 |||||
 Db 1646 TGTCTCTGAGTCTTGAAGCACTGAGCTGTGCTGAGAGCTCTTCCGAAAGCCCACTGC 1705
 |||||
 QY 1727 GGGGCTGTGTGTCTGAGCAACTGAGGAGGCTTGAAGCACTCTCTCTGAGAG----- 1780
 |||||
 Db 1706 GGGGCTGTGTGTCTGAGCAACTGAGGAGGCTTGAAGAGCCCTCTGAGAGGAGCTCTG 1765
 |||||
 QY 1781 -----GTCTGAAGCGATGTGAGAGATCTTGGGGAGTAGCTGTTTCTCTGACAG 1834
 |||||
 Db 1766 GAAGCTGCTGAGGAGGCTGAGGAGGATCTTGGGGAGCAGCTGTTTCTCTGACAG 1825
 |||||
 QY 1835 ACTGCCAAGAGTGAAGCTGAGCTTACAGACAGAGCCCAAGAACTGTCTCAAGCTCAGC 1893
 |||||
 Db 1826 ACTGCCAAGAGTGAAGCTGAGCTTACAGACAGAGCTGAGGAGTGTCTCAAGCTCAGC 1884
 |||||
 RESULT 5
 ABK14000
 ID ABK14000 standard; cDNA; 3353 BP.
 AC ABK14000;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE CDNA encoding human protein kinase 3700.
 XX
 XX Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
 KM cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
 KM tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
 KM cell proliferation disorder; cell differentiation disorder; carcinoma;
 KM haemotopoietic neoplastic disorder; metastatic disorder; leukemia;
 KM cytoskeletal; antithrombotic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT 157..2043
 FT CDS
 FT /product= "Protein kinase 3700"
 FT /note= "The coding region (not including the terminator

FT PT (codon) is specifically claimed in claims 7 and 28"
 XX
 XX WO200224921-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US030115.
 XX
 PR 25-SEP-2000; 2000US-0234922P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ, Galvin KM;
 XX
 DR WPI; 2002-352007/38.
 DR P-PDB; AAU79652.
 DR
 XX
 XX Use of modulators of activity of 3700 protein for making medicament for
 PT e.g., modulating protein phosphorylation or cell signalling, or for
 PT treating or preventing cellular proliferative and/or differentiative
 PT disorders.
 XX
 PS Claim 28; Fig 1; 115bp; English.
 XX
 CC The present invention relates to the isolation of a novel human protein
 CC kinase designated 3700, and the polynucleotide sequence encoding it. The
 CC invention also describes the use of a modulator of the activity of
 CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
 CC composition for modulating the ability of a cell to phosphorylate an
 CC amino acid residue of a substrate protein. Modulators of protein kinase
 CC 3700 activity are useful for modulating protein phosphorylation, cell
 CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
 CC angiogenesis, tissue repair, tissue regeneration, establishment or
 CC progression of atherosclerosis, and signalling across the blood-brain
 CC barrier. The polynucleotide and polypeptide molecules for protein kinase
 CC 3700 may be used as diagnostic targets and therapeutic agents for
 CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
 CC curing PK-related disorders and cellular proliferative and/or
 CC differentiative disorders (e.g. haematopoietic neoplastic disorders,
 CC carcinoma, sarcoma, metastatic disorders or leukemia). The 3700
 CC polynucleotide sequence can be used to express protein kinase 3700, to
 CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for
 CC tissue typing, in forensic biology, and as surrogate markers. The present
 CC sequence encodes human protein kinase 3700
 XX
 SQ Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;
 Query Match 71.9%; Score 1360.2; DB 6; Length 3353;
 Best Local Similarity 84.2%; Pred. No. 0;
 Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;
 QY 47 CTTCCGCTTGGCTGAGAGCGCCGCGCTGCGGAGCGGGCTCATCAATCGCTTA 106
 |||||
 Db 191 CCACTCCCTCGCGCGAGCTAGCCCGCGCTGCGGAGGAGGCTGATCAATCGGCCA 250
 |||||
 QY 107 AACCTCGATGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 166
 |||||
 Db 251 AGCCCTTAATGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 310
 |||||
 QY 167 GCTACGAGTTCCTGAGAGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 226
 |||||
 Db 311 GCTACGAGTTCCTGAGAGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 370
 |||||
 QY 227 AGAGCTCGGGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 286
 |||||
 Db 371 AGAGCTCGGGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 430
 |||||
 QY 287 AGGATCGCTGAGCAGATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 346
 |||||
 Db 431 AAGATCTGATGACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 490
 |||||
 QY 347 TCATTGCATCATGAAGTGTGAGAAATGAGCAAGATGTGATGTGATGAGGATGATG 406
 |||||

Db 491 TCATTGCCATCATGAAGTGTGAGAACAGCAGCAAGATCGTATGCTCATGAGATATG 550
QY 407 CCAGCCGAGGCGATCTGTATGATTAATCATGAGAGCGCCACCGCTGATGAGCGGAGC 466
Db 551 CCAGCCGAGGCGATCTGTATGATTAATCATGAGAGCGCCACCGCTGATGAGCGGAGC 610
QY 467 CCAGGAGATTTCTTCCGACAGATGCTGTGCGCCGCACTACTGCAACAGAGGATGCG 526
Db 611 CTAGGATTTCTTCCGACAGATGCTGTGCGCCGCACTACTGCAACAGAGGATGCG 670
QY 527 TTCACCGAGATCTCAAGCTGAAAACATCTCTTAGATGCAATGAAACATCAAGATTTG 586
Db 671 TCACCGAGATCTCAAGCTGAAAACATCTCTTAGATGCAATGAAACATCAAGATTTG 730
QY 587 CTGACTTTGGCCTCTTCCACCTGTATCAACAAAGCAAGTTCTTCCAGACCTTCTGTGGA 646
Db 731 CTGACTTTGGCCTCTTCCACCTGTATCAACAAAGCAAGTTCTTCCAGACCTTCTGTGGA 790
QY 647 GCCCTCTTACGCGCTGCGCTGAGATAGTCAACGAGGAGCCCTATGATGAGCGAGGTTG 706
Db 791 GCCCTCTTACGCGCTGCGCGAGATAGTCAACGAGGAGCCCTATCAACGAGCGAGGTTG 850
QY 707 ACAGCTGTCTCTTGGGCGTTCTCTGTATCATCTGTGTATGCAACCATGCTTGTAGC 766
Db 851 ACAGCTGTCTCTTGGGCGTTCTCTGTATCATCTGTGTATGCAACCATGCTTGTAGC 910
QY 767 GGGAGATCATTAACACTGTGTAAAGAAATCAATGAAACGAGGCTTACCTGTAGCGGCCA 826
Db 911 GGGATACCAATTAAGATCTTGTGTAAACAGATCAAGCAACGAGGCTTACCTGTAGCGGCCA 970
QY 827 AGCGCTCCATGCTGTGCGCTGATTCGGGCGCTGTTAATGTTGAACCCACCTGTCGGG 886
Db 971 AACCTCTGATGCTGTGCGCTGATTCGGGCGCTGTTAATGTTGAACCCACCTGTCGGG 1030
QY 887 CCACACTGAGAGATGATAGCAGATCATTTGTGGGTCAACTGGGTTATACACCGAGTGC 946
Db 1031 CCACCGTGAAGATGATGAGGCACTGTGGGTCAACTGGGTTATACACCGAGTGC 1090
QY 947 GGGAAACAGAAAGCCCTGCGTGAAGGCTGAGGCAACCTTGTGTGACTTTGGCGGCTTCA 1006
Db 1091 GAGAGACAGAGGCTGCGCATGAGGCTGAGGCAACCTTGTGTGACTTTGGCGGCTTCA 1150
QY 1007 TGCGGAGTGTGTTACGTGCTCTCGCGCCCTCTCTGAGATGATGAGCCCAAGGTGTGCA 1066
Db 1151 TGCGTACTGTGCTCGCGGCTTCTCGCGCCCTCTCTGAGATGATGAGGCTTCAAGGTGTGCA 1210
QY 1067 GCTTCTTCAAGGACAGCTGCGCGAGGTGAAAGCACTTACTTGGGCTGAGCGGCAAC 1126
Db 1211 GCTTCTTCAAGGACAGCTGCGCGAGGTGAAAGCACTTACTTGGGCTGAGCGGCAAC 1270
QY 1127 ATTCTCTTAAGAAAGTCCGAAAGAGATGATGCTCAAAATCTGCAAGGTGACCCGG 1186
Db 1271 ATTCTCTTAAGAAAGTCCGCGAAGAGATGATGCTCAAAATCTGCAAGGTGACCCGG 1330
QY 1187 CTGAGGATACCTCTTCTCGCGCTGCAAGGACGCTTAAAGCTTCCGAAGGCAATTTCTCA 1246
Db 1331 CTGATTAACCTGCGCAATGCGCTGCGAAGGCAACCTTCAAGCTTCCGAAGGCAATTTCTCA 1390
QY 1247 AGAAAAAGTCTTAACTACCTGTCAGGAGAGGTATCAGAGAGACCTTCAAGAACTCAGACCGG 1306
Db 1391 AGAAGAGAGGTATGAGCTTCTGCAAGAAAGGTTATAGAGAGACCTTCCGAGCTTCAAGCCCA 1450
QY 1307 TGCCTGATTAATCTCAGAGGAGACCTGTGCTGCTGATTCCTGCTCCAAAGAAAGGACATCC 1366
Db 1451 TGCCTGAGAGCCAGAGGAGAGGCTGCGC-----CCCTGCTCCCAAGAAAGGACATTC 1501
QY 1367 TTAAGAGAGTCTGACAGCGGTGATCTGGTTACTTACTTCTTCCAGAGCCAGAGATCTTG 1426
Db 1502 TCAGAGAGCCCGACAGCGGAGTCTGGCTACTTCTCTTCCAGAGCCAGAGATCTTG 1561
QY 1427 GGGAACTCTTGAAGCGCAGATGATGTTTGTGATGAGGAGCCCGGAGAGCAGAAATCTTC 1486
Db 1562 GGGAGCTCTTGAAGCGAGGAGAGCTGTTTGTGATGAGGAGATCCCAAGAGACAGAGCTTC 1621

QY 1487 CACAGCTTCAGGCTTCCTCTCCACCGAAGGCAATTCATAATGCAAGTTCT 1546
Db 1622 CGCAAGCTTCAGGCTTCCTCTCCATGCAAGGCAATCTCAATGCAAGTTCT 1681
QY 1547 CCGCAGAGCTTAAAGAGCACTAACCCCTTAGACACTTTGGCTCCCTGACCACTGGCT 1606
Db 1682 CCGAGAGCTTAAAGAGCTTACCGGCCCCCAACCTTGGCTCCCTGAGAACTCGCCC 1741
QY 1607 CCTCCATCTGACAGCGCGGCCCGACCGCCCTCAGGAGGCTGTAGAGAGAGCATCC 1666
Db 1742 CACCTGCGCCCTGCGCCCGGCGAGCCGACCTCAGAGGCTGTAGAGAGAGCATCC 1801
QY 1667 TGTCTTCAGATCCTTGAACCAATGGAATTTGCTGAAACGTTCCGAAACCCCACTGA 1726
Db 1802 TGTCTTCAGATCCTTGAACCACTGGAATTTGCTGAAACGTTCCGAAACCCCACTGC 1861
QY 1727 GGGGCTGTGTGTGTGTGAGCAACCTGAGGAGGCTTGAAGAGCTCCCTCAGAAAG----- 1780
Db 1862 GGGGCTGTGTGTGTGTGAGCAACCTGAGGAGGCTTGAAGAGGCTCCCTCAGAGGAGCTGTG 1821
QY 1781 -----GTCTGAAGCGATGAGTGGCAGAGAAATCTTGGGAGATGATGCTGCTTCTGTAG 1834
Db 1922 GAAGCTGCTGAGGCGCTGCGGAGAGATCTTGGGAGACAGCTGCTTTTCCCTGAGAG 1981
QY 1835 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAACAGCCCTAGGAATCTCTCAAGCTCAGC 1893
Db 1982 ACTGCCAAGAGTGAAGTGAAGCTTGAAGCAACAGCTTGAAGGAGCTCTCAAGCTCAGC 2040

RESULT 6
ADL14160
ID ADL14160 standard; cDNA; 3353 BP.
XX
AC ADL14160;
XX
DT 17-JUN-2004 (first entry)
XX
DE Novel human gene 3700 cDNA.
XX
KW cytosolic; cardiac; hypotensive; antianginal; osteoporotic;
KW antirheumatic; antirheumatic; neuroprotective; antiinflammatory;
KW antidiabetic; antidiabetic; cardiovascular; virucide; analgesic; CNS;
KW angiotensin inhibitor; angiotensin stimulator; cerebroprotective;
KW nephrotropic; antihypertensive; dermatological; immunomodulator;
KW cell proliferation disorder; cell differentiation disorder;
KW brain disorder; platelet disorder; breast disorder; colon disorder;
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;
KW thyroid disorder; testes disorder; haematopoietic disorder;
KW pancreatic disorder; skeletal muscle disorder; skin disorder;
KW dermal disorder; bone metabolism disorder; immune disorder;
KW inflammatory disorder; cardiovascular disorder;
KW endometrial cell disorder; liver disorder; viral disease; pain disorder;
KW metabolic disorder; neurological disorder;
KW central nervous system disorder; erythroid disorder;
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
KW cell proliferation; cell differentiation; cell growth; cell division;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN US2004058355-A1.
XX
PD 25-MAR-2004.
XX
PF 25-APR-2003; 2003US-00423543.
XX
PR 30-SEP-1998; 98US-00163821.
PR 27-JUN-1999; 99US-0117580P.
PR 25-MAR-1999; 99US-0027640P.

CC hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity
CC of a test compound and in gene therapy. The present sequence is human
CC PKIN-12 cDNA
XX

SO Sequence 3360 BP; 729 A; 1022 C; 882 G; 727 T; 0 U; 0 Other;

Query Match 71.7%; Score 1357.6; DB 6; Length 3360;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;

```
QY 32 GCCAGGTCCTCCGCGCTCCGCTCCGCTCCGAGAGAGCCCGCCCGCTGCGAGCGGC 91
DB 98 GCGGGGCTCCGCGCTCCGCTCCGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 157
QY 92 TCATCAAGTCGCTTAACTCTGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151
DB 158 TGATCAAGTCGCTTAACTCTGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 217
QY 152 ACAACCTGCGGCAACCGCTACGAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 211
DB 218 ACAACCTGCGGCAACCGCTACGAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 277
QY 212 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 271
DB 278 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 337
QY 272 AAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
DB 338 AAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397
QY 332 TCAACCAACCCCAATCATTTGCCATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 391
DB 398 TCAACCAACCCCAATCATTTGCCATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
QY 392 TTGTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451
DB 458 TTGTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
QY 452 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511
DB 518 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
QY 512 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 571
DB 578 ATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
QY 572 GAAACATCAAGATGCTGACTTGGCTCTTCAACCTGTACCAAGAGAGAGAGAGAGAG 631
DB 638 GAAATATCAAGATGCTGACTTGGCTCTTCAACCTGTACCAAGAGAGAGAGAGAGAG 697
QY 632 AGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 691
DB 698 AGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
QY 692 TGGGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
DB 758 CAGGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
QY 752 CCAATGCCCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
DB 818 CCAATGCCCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 877
QY 812 ACCGTGAGCGGCGCAAGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871
DB 878 ACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
QY 872 ACCCCACCGGTGCGGCGCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931
DB 938 ACCCCACCGGTGCGGCGCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 997
QY 932 ACACCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 991
DB 998 ACACCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1057
```

```
QY 992 TTGAGCGAGGCTCTCATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1051
DB 1058 CTGCGCGGCTCTCATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
QY 1052 GAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111
DB 1118 GAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
QY 1112 GGCCTGAGCGGCAACATTCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
DB 1178 GGCCTGAGCGGCAACATTCCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
QY 1172 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231
DB 1238 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
QY 1232 GAAAGGCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1291
DB 1298 GAAAGGCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
QY 1292 AGGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1351
DB 1358 AGGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
QY 1352 CAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411
DB 1409 CAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1468
QY 1412 AGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1471
DB 1469 AGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528
QY 1472 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531
DB 1529 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1588
QY 1532 TCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591
DB 1589 TCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1648
QY 1592 TGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1651
DB 1649 TGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1708
QY 1652 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1711
DB 1709 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1768
QY 1712 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
DB 1769 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1828
QY 1772 CCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
DB 1829 CCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1888
QY 1820 GCTTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1879
DB 1889 GCTTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1948
QY 1880 GCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1893
DB 1949 GCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1962
```

RESULT 8
AAK94280
ID AAK94280 standard; cDNA; 3395 BP.
XX AAK94280;
AC
XX
DT 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2918.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX (HELI-) HELIX RES INST.
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
XX Makamatsu A, Sugiyama T, Negai K, Kojima S, Otsuki T, Koga H;
XX WPI, 2001-524255/58.
XX P-PSDB; AAM93360.
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX in genetic manipulation.
XX Claim 8; SEQ ID NO 2918; 1380bp + Sequence listing; English.
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesized by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a full length human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in CD-ROM format directly
XX from EPO
XX Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
Query Match 71.7%; Score 1357; DB 4; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCGCGCCCTGGCGCTGGAGAGCGCCCGCGCTGGCGGACGGCTCATCAAGTCCCTTA 106
DB 152 CCACTCCCTCGCGCGAGAGCTAGCCCGCGCTGGCGGAGGGCTGATCAATCCGCCA 211
QY 107 AACCTCTGTAAGAAAGACAGCGCGGTGAAGCGGCACCATCAAAACAACCTCGGACCC 166
DB 212 ACCCCCTATGAAGAAAGAGCGCGGTGAAGCGGCACCAACAGCAACACTCGGACCC 271
QY 167 GCTAGAGGTTCTCGAGAGCGTGGGCAAGGGCACTTACGGGAAGTGAAGAAAGCAAG 226
DB 272 GCTACAGATTCTCGAGAGCGTGGGCAAGGGCACTTACGGGAAGTGAAGAAAGCGCG 331
QY 227 AAGAGCTCGGGGGGTCTGGTGGCCATCAATGCTCATGAGAAAGCAAAATCAAGATGAGC 286
DB 332 AAGAGCTCGGGGGGTCTGGTGGCCATCAATGCTCATGAGAAAGCAAAATCAAGATGAGC 391
QY 287 AAGATCTGTCGACATACGAGAGGAGATGAGATGATGCTTCACTCAACCAACCCACACA 346
DB 392 AAGATCTGTCGACATACGAGAGGAGATGAGATGATGCTTCACTCAACCAACCCACACA 451
QY 347 TCATTGCCATCAATGAAGTGTGTTGAGAAATGACAGCAAGATGTGATGTTCATGAGTATG 406
DB 452 TCATTGCCATCAATGAAGTGTGTTGAGAAATGACAGCAAGATGTGATGTTCATGAGTATG 511

QY 407 CCAGCCGAGGCGATCTGTATGATTACATCAATGAGGGGCCACCGGCTGATGAGCGGACG 466
DB 512 CCAGCCGAGGCGATCTGTATGATTACATCAATGAGGGGCCACCGGCTGATGAGCGGACG 571
QY 467 CCAGGATTTCTTCCGACAGATGCTGTGCTCCCTGACACTGCTCCACAGAGGGATGCG 526
DB 572 CTAGGATTTCTTCCGAGAGATGCTGTGCTCCCTGACACTGCTCCACAGAGGGATGCG 631
QY 527 TTCACGAGATCTCAAGCTGGAAGAAACATCTTCAATGCGAATGGAACATCAAGATTG 586
DB 632 TCCACGAGATCTCAAGCTGGAAGAAACATCTTCAATGCGAATGGAACATCAAGATTG 691
QY 587 CTGACTTGGCCCTTCCAACTCTTACCAACAAAGCAAGTTCTCCAGAGCTTCTGTGGGA 646
DB 692 CTGACTTGGCCCTTCCAACTCTTACCAACAAAGCAAGTTCTCCAGAGCTTCTGTGGGA 751
QY 647 GCCCTCTACGCGCTGCGCTGAGATGTCACAGGGAAGCCCTATGAGGGCCAGAGGTGG 706
DB 752 GCCCTCTACGCGCTGCGCTGAGATGTCACAGGGAAGCCCTATGAGGGCCAGAGGTGG 811
QY 707 ACAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGACATGCGACATGCGCTTGTAGC 766
DB 812 ACAGCTGTCTCTGGGCGTTCTCTGTACATCTGTGTGACATGCGACATGCGCTTGTAGC 871
QY 767 GCGAGGATCATTAATAACACTGTGGAAGCAATCAATCAAGGCGTTACCGTGAGCCGCCA 826
DB 872 GCGATACCATTAAGATCTTAGTGAACAATCAAGCAAGCGGCGTTACCGTGAGCCGCCA 931
QY 827 AGCGCTCCAGTGCCTGTGCGCTGATCCGGTGGCTGTTAATGGAAGCCCAACCGTGGG 886
DB 932 AACCTCTGATGCTGTGCGCTGATCCGGTGGCTGTTAATGGAAGCCCAACCGTGGG 991
QY 887 CCACTGTGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 992 CCACTGTGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1051
QY 947 GGAAGAAGAAAGCGCTGCGTGAAGGAGGAGCACTTGTGATGATGATGATGATGATGATG 1006
DB 1052 GGAAGAAGAAAGCGCTGCGTGAAGGAGGAGCACTTGTGATGATGATGATGATGATGATG 1111
QY 1007 TGGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 1112 TGGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171
QY 1067 GCTTCTTCAAGAGCAAGTGGCGGAGTGAAGCACTGATCTGAGCTGAGCGGCAAC 1126
DB 1172 GCTTCTTCAAGAGCAAGTGGCGGAGTGAAGCACTGATCTGAGCTGAGCGGCAAC 1231
QY 1127 ATTCTTGAAGAAAGTCCGAAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCG 1186
DB 1232 ATTCTTGAAGAAAGTCCGAAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCG 1291
QY 1187 CTGAGGATACCTTCTTGGCCCTGCAAGAGCAAGCTTAAAGTCTCGGAAGGATCTCA 1246
DB 1292 CTGATGATCTGCTTCCATTCGCTTGGCAAGAGCACTTAAAGTCTCGGAAGGATCTCA 1351
QY 1247 AAAAAAAGTCTTCACTTCTGTCAGGAGAGTGAAGAGGACCTTCAAGAACTCAGACCG 1306
DB 1352 AAAAAAAGTCTTCACTTCTGTCAGGAGAGTGAAGAGGACCTTCAAGAACTCAGACCG 1411
QY 1307 TGGCTGATCTCAAGGAGAGCTGTCCTGCTGTATCTCTGCTCCCAAGAAAGGATCC 1366
DB 1412 TGGCTGATCTCAAGGAGAGCTGTCCTGCTGTATCTCTGCTCCCAAGAAAGGATCC 1462
QY 1367 TTAAGAAAGTCTGACAGAGGTAATGCTGTTACTACCTCTCAAGAGCCCAAGAGTCTG 1426
DB 1463 TTAAGAAAGTCTGACAGAGGTAATGCTGTTACTACCTCTCAAGAGCCCAAGAGTCTG 1522
QY 1427 GGAAGCTTGAAGAGCAAGTGAATGTTGTTGAGTGGGAGCCCGTGAAGCAAGTCTC 1486
DB 1523 GGAAGCTTGAAGAGCAAGTGAATGTTGTTGAGTGGGAGCCCGTGAAGCAAGTCTC 1582
QY 1487 CAAGGCTTCAAGGCTCTCTCTCAACCGCAAGGCAATCTCAAACTCAATGCAAGTCT 1546

Db 1583 CGAAGCTTCAGGAGCTGCTCTCCATGCAAGCATCTCAATGAGCAAGTTCT 1642
Qy 1547 CCCGACAGGCTTGAAGAGCACTACCCCTAGCACTTTGGCTCCCTGACCAATGGGCT 1606
Db 1643 CCCAGACAGCTTGGAGCTGCGGCCCCCAGACACTTGGCTCCCTGAGAACTGGCCC 1702
Qy 1607 CCTCCATCTCTGAGCCGCGCCAGCCGCTCAGGGGCTGTGATGAGAGACAGATCC 1666
Db 1703 CACCTGGCCCCCTGGGCCCGGAGCGACCTCAGGGGCTGTGAGAGAGAGATCC 1762
Qy 1667 TGTCTCTGAGCTCTTGAACCAATTGGAATTGCTGCTGAAGCTCTTCCGAAACCTCACTGA 1726
Db 1763 TGTCTCTGAGCTCTTGAACCAAGCTGCACTGCTGCAAGCGCTCCAGAGCCCCCACTGC 1822
Qy 1727 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCCCTCCAGAAAG----- 1780
Db 1823 GGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCCCTCCAGAGGGCCCTG 1882
Qy 1781 -----GTCTGAAGCAGATGTGAGAGGAATCTTGGGGATAGCTGCTTTCTGTGACAG 1834
Db 1883 GAAGCTGCTGAGGCGCTGCGGAGATCTTTGGGGAGACGCTGCTTTCCCTGACAG 1942
Qy 1835 ACTGCCAAGAGTGACTGACGCTTACAGACAGAGCCCTAGAGATCTGCTCAAGCTCAGC 1893
Db 1943 ACTGCCAAGAGTGACAGGACCTTACCGACAGGCACTGAGGGTCTGCTCAAGCTCAGC 2001

RESULT 9

ADL30885 ID ADL30885 standard; cDNA; 3395 BP.

XX AC ADL30885;

XX DT 20-MAY-2004 (first entry)

XX DE Full length human cDNA clone Segid 2918.

XX human; medicine; signal transduction; glycoprotein; transcription;

XX oligo-capping method; 88; gene.

XX OS Homo sapiens.

XX PN EP1396543-A2.

XX PD 10-MAR-2004.

XX PF 07-JUL-2000; 2003BP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JUN-2000; 2000UP-00118774.

XX PR 02-MAY-2000; 2000UP-00183865.

XX PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI; 2004-204755/20.

XX DR P-PSDB; ADL30886.

XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX length human cDNAs.

XX Example 1; SEQ ID NO 2918; 1340BP; English.

XX This invention relates to a novel primers useful for synthesizing full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction,

XX glycoproteins and transcription. The present invention describes a method

CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.

SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;

Query Match 71.7%; Score 1357; DB 12; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

Qy 47 CTTCCGCTGCTGCTGAGAGCGCCGCGCTGCGGAGCGGCTCATGAAGTGGCTTA 106
Db 152 CCACTCCCTCGCGCGAGACTAGCCCGGCTGCGGAGAGGCTGATGAAGTGGCTTA 211
Qy 107 AACCTGATGAAG 166
Db 212 AGCCCTGATGAAG 271
Qy 167 GCTACAGATTCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 226
Db 272 GTACAGATTCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
Qy 227 AGAGCTGAG 286
Db 332 AGAGCTGAG 391
Qy 287 AGAGCTGAG 346
Db 392 AGAGCTGAG 451
Qy 347 TCATTGCTGATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
Db 452 TCATTGCTGATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511
Qy 407 CCAAGCGAG 466
Db 512 CCAAGCGAG 571
Qy 467 CCAAGCGAG 526
Db 572 CCAAGCGAG 631
Qy 527 TTCACGAG 586
Db 632 TTCACGAG 691
Qy 587 CTGACTTGGCTCTGCAACCTGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646
Db 692 CTGACTTGGCTCTGCAACCTGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
Qy 647 GCGCTCTGACGCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
Db 752 GCGCTCTGACGCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811
Qy 707 ACAAGCTGCTGAG 766
Db 812 ACAAGCTGCTGAG 871
Qy 767 GGCAGAGATCAATAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
Db 872 GGCAGAGATCAATAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 931
Qy 827 AGCCCTGATGAG 886
Db 932 AGCCCTGATGAG 991
Qy 887 CCAAGCTGAG 946
Db 992 CCAAGCTGAG 1051
Qy 947 GGAAG 1006
Db 1052 GGAAG 1111

QY 1007 TGGGAGACGTGGTTACGTGCGCTCGCGCCCTCCCTGGAAGAAATGAGCCGAAGTGTGCA 1066
DB 1112 TGGCTGACGTGGCTCGCGGCTTCCTCCGCCCTCTGAGAAATGGGCGCAAGGTGTGCA 1171
QY 1067 GCTTCTTCAAGCAGCAGCTGCGCGGAGTGTGAAACACTGTAACCTGGCTGAGCGGCAAC 1126
DB 1172 GCTTCTTCAAGCAGCAGCTGCTGTGGGGAAAGCAACACCCCTGGCTGTGAGGCGCAGC 1231
QY 1127 ATTCTCTTAAGAAAGTCCCGAAAGGAATGACATGCTCAAAATCTGCAAGTGTGACCCGG 1186
DB 1232 ATTGCTCAAGAAATGCTCCGCAAGGAATGACATGAGCCCACTCTCTCAAGTGTGACCGG 1291
QY 1187 CTGAGGATACCTCTCTCGCTCGCGCAAGAGCCCTTAAGCTTCCGAAAGGCAATTCCTCA 1246
DB 1292 CTGATGACCTGCTCCATCGCTCGCAAGAGCACTCAAGCTGCAAGGCGCAATTCCTCA 1351
QY 1247 AGAAAGATCTCTTACCTGCTCAAGGAGGTGACAGAGAGACCCCTCAGAACTCAGACCGG 1306
DB 1352 AGAAGAAAGTGTGACCTCTGCAAGAAAGGAGTACAGAGAGACCTCCGAGACTCAGCCCA 1411
QY 1307 TGCCTGATCTCAAGGAGAGCTGTCTGCTGTATCTCTGCTCCCAAGAAAGGCAATCC 1366
DB 1412 TCCCTGAGGCCCAAGGAGGAGGCTGCCC-----CCCTGCTCCCAAGAAAGGCAATTC 1462
QY 1367 TTAAGAGTCTGACAGCGGTGAATCTGGTTACTACTCTCTCCAGAGCCCAAGGAGCTTG 1426
DB 1463 TCAAGAGGCCCAAGGAGGAGTGTGCTGCTACTCTCTCTCCGAGCCCAAGTGAATCTG 1522
QY 1427 GGGAACTCTTGAAGCGCAGTGTATGTTGTGTGAGTGGGAGACCCCGTGAAGAGTCTC 1486
DB 1523 GGGAGCTCTTGAAGCGAGGAGAGTGTGTGTGAGTGGAGATCCCAAGAGACAGAGCTC 1582
QY 1487 CAGAGCTTGAAGGAGCTCTCTCTCAACGCAAGGAGCTTCAAACTCAATGAGCAAGTCT 1546
DB 1583 CGAGAGCTTGAAGGAGCTCTCTCTCAAGGAGCACTCAAACTCAATGAGCAAGTCT 1642
QY 1547 CCGCAGAGCTTGAAGAGCACTACCCCTGAGCACTTGTGCTCCCTGAGCACTAGGCT 1606
DB 1643 CCGAGAGCTTGAAGAGCTGCGAGCCCAACACTTGTGCTCCCTGAGCACTAGGCT 1702
QY 1607 CCTCCATCTGAGCGCGGCGCAGCGCCCTCAGGAGGCTGTGAGTGAAGAGCACTCC 1666
DB 1703 CACTGCGCCCGCGCGCGCGGCGCAGCGCACTCAGGAGGCTGTGAGGAGAGAGCACTCC 1762
QY 1667 TGTCTCTCCAGTCTTGTGACCAATTGACATCTTGCTGCTTCCGAAACCCCACTGA 1726
DB 1763 TGTCTCTGAGTCTTGTGACCACTGGAATCTTGCTGGAACGAGTCCCGAGACCCCACTGC 1822
QY 1727 GGGGCTGTGTGTGTGTGAGCACTGAGGGGCTTGAAGAGCTCCCTCAGAA----- 1780
DB 1823 GGGGCTGTGTGTGTGTGAGCACTGAGGGGCTTGAAGAGCCCTCTCAGAGGGCCCTG 1882
QY 1781 -----GTCTGAAGCAATGTGTGAGCAATCTTGGGGAGTAGCTGTTTCTCTGAG 1834
DB 1883 GAAAGCTGCTGAGGCGCTGCGCGCAGAGATCTTTGGGGAGCAGCTGTTTCCCTGAG 1942
QY 1835 ACTGCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1893
DB 1943 ACTGCCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2001

RESULT 10
ABD33085
ID ABD33085 standard; cDNA; 3404 BP.
XX
AC ABD33085;
XX
XX 18-NOV-2004 (first entry)
XX
DE Human cancer-associated (CA) cDNA HR07-004.
XX
XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW

KW 88; cancer; cyrostatic.
OS Homo sapiens.
PN WO2004058146-A2.
PD 15-JUL-2004.
PF 15-DEC-2003; 2003WO-US040081.
PR 17-DEC-2002; 2002US-00322281.
PA (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-499109/47.
DR P-PSDB; AB084757.
XX
PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Claim 1; SEQ ID NO 25; 182pp; English.
XX
CC The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents human CA cDNA of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences
XX
SQ Sequence 3404 BP; 723 A; 1044 C; 895 G; 742 T; 0 U; 0 Other;
Query Match 71.7%; Score 1357; DB 13; Length 3404;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCGCGCCCGCTGCGAGAGCGCGCGCGCTGCGGAGCGGCTCATCAAGTGGCTTA 106
DB 162 CCACTCCCTCGGCGGAGAGTACCGGCGCGCTGCGGAGCGGCTCATCAAGTGGCTTA 221
QY 107 AACCTGTATGAAGAGCAGCGGCTGAGCGGCACTATCAAAACAACCTGCGGAC 166
DB 222 AGCCCTTAATGAAGAGCAGCGGCTGAGCGGCACTATCAAAACAACCTGCGGAC 281
QY 167 GCTACGAGTCTCTGAGAGCGCTGCGGAGCGCACTTACCGGAAAGTGAAGAGCAGAG 226
DB 282 GCTACGAGTCTCTGAGAGCGCTGCGGAGCGCACTTACCGGAAAGTGAAGAGCAGAG 341
QY 227 AGAGCTCGGGGGGCTGCGGAGCGCATCAAGTCAAGTGAAGAAACAACCAAGAGAGAC 286
DB 342 AGAGCTCGGGGGGCTGCGGAGCGCATCAAGTCAAGTGAAGAAACAACCAAGAGAGAC 401
QY 287 AGAGCTGCTGACATACGAGGAGAGATGAGATCATGCTTCACTCAACACCCGAC 346
DB 402 AGAGTGTATGACATACGAGGAGAGATGAGATCATGCTTCACTCAACACCCGAC 461
QY 347 TCATTGCCATCCATGAAGTGTGAGAAATGAGCAGCAAGATGTGATGTGATGAGATG 406
DB 462 TCATTGCCATCCATGAAGTGTGAGAAATGAGCAGCAAGATGTGATGTGATGAGATG 521
QY 407 CCAAGCGAGCGCATCTGTATGATTAATCAAGTGAAGGCGCAGCGCTGATGAGCGGAGG 466

Db 522 CAGAGCGGGGCACTTTATGACTACATCAAGCGAGCGGACAGCTCATGAGCGCGAG 581
 Qy 467 CCAGGAGATTCTTCCGACAGATGATGTCGCTCCCTGCACTACTGCAACAGAGCGGATCG 526
 Db 582 CTAGGAGATTCTTCCGAGAGATGATGTCGCTCCCTGCACTACTGCAACAGAGCGGATCG 641
 Qy 527 TTCAAGGAGATCTCAAGAGTGGAAAACATCTTTCTAGATGCGAATGGAACATCAAGATTG 586
 Db 642 TCACAGAGATCTCAAGAGTGGAAAACATCTTTCTAGATGCGAATGGAACATCAAGATTG 701
 Qy 587 CTAGATTTGGGCTCTCCAACTGATCAACAGAGCGGATCTTCCAGAGCTTCTGAGGGA 646
 Db 702 CTAGATTTGGGCTCTCCAACTGATCAACAGAGCGGATCTTCCAGAGCTTCTGAGGGA 761
 Qy 647 GCGCTCTCTAGCGCTCGCTGAGATAGTCAACGAGGAGCGCTTATGTTGGGCTCAGAGGTTG 706
 Db 762 GCGCTCTCTAGCGCTCGCTGAGATAGTCAACGAGGAGCGCTTATGTTGGGCTCAGAGGTTG 821
 Qy 707 ACAGCTGATCTGAGGCGTTCTCTGTACATCTCTGTGATGAGCAACGATGCGCTTTGACG 766
 Db 822 ACAGCTGATCTGAGGCGTTCTCTGTACATCTCTGTGATGAGCAACGATGCGCTTTGACG 881
 Qy 767 GCGAGATCTATAAACAATGATGAGGAATCAATGAAACGAGGCTTACCTGAGGCGGCA 826
 Db 882 GCGATGACATTAAGATCTTAGTAAACAATGAAACAAGGAGCGCTTACCTGAGGCGGCA 941
 Qy 827 AGCGCTCCGATGCTGCTGAGCTGATCCGAGGCTGATTAATGATGAAACCCGACCGGCTGAG 886
 Db 942 AACCTCTGATGCTGCTGAGCTGATCCGAGGCTGATTAATGATGAAACCCGACCGGCTGAG 1001
 Qy 887 CCACTGAGAGATGATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
 Db 1002 CCACTGAGAGATGATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1061
 Qy 947 GGGAAACAGAGAGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1006
 Db 1062 GAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1121
 Qy 1007 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1066
 Db 1122 TGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1181
 Qy 1067 GCTTCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126
 Db 1182 GCTTCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1241
 Qy 1127 ATTCTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1186
 Db 1242 ATTCTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1301
 Qy 1187 CTAGAGATGATCTTCTGCGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1246
 Db 1302 CTAGATGATCTTCTGCGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1361
 Qy 1247 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1306
 Db 1362 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1421
 Qy 1307 TGCCTGATCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1366
 Db 1422 TGCCTGATCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1472
 Qy 1367 TTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1426
 Db 1473 TCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1532
 Qy 1427 GGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1486
 Db 1533 GGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592
 Qy 1487 CAGAG 1546
 Db 1593 CAGAG 1652

Qy 1547 CCGCAGAGGCTTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1606
 Db 1653 CCGCAGAGGCTTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1712
 Qy 1607 CTTCCAGATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1666
 Db 1713 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1772
 Qy 1667 TGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1726
 Db 1773 TGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1832
 Qy 1727 GGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1780
 Db 1833 GGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1892
 Qy 1781 -----GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1834
 Db 1893 GAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1952
 Qy 1835 ACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1893
 Db 1953 ACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2011

RESULT 11
 ADJ96554
 ID ADJ96554 standard; DNA; 3463 BP.
 XX
 AC ADJ96554;
 XX
 PT 06-MAY-2004 (first entry)
 XX
 DE Human calcium/calmodulin-dependent protein kinase Nnak2 DNA Seqid 11.
 XX
 XX gene; de; kinase; human; SNP; single nucleotide polymorphism;
 KW tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
 KW gene therapy; cancer; immune-related disease; cardiovascular disease;
 KW brain; neuronal associated disease; metabolic; inflammatory disorder;
 KW cytoskeletal; neuroprotective; immunomodulator; antiinflammatory;
 KW calcium/calmodulin-dependent protein kinase; Nnak2.
 XX
 OS Homo sapiens.
 OS 68.
 XX
 XX Key Location/Qualifiers
 FH variation replace(1670,c)
 FT /*tag= a
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(1727,a)
 FT /*tag= b
 FT /standard_name= "Single nucleotide polymorphism"
 XX
 XX MO2004006838-A2.
 XX
 XX PD 22-JAN-2004.
 XX
 XX PF 15-JUL-2003; 2003MO-US021730.
 XX
 XX PR 15-JUL-2002; 2002US-0395632P.
 XX
 XX (SUGF-) SUGEN INC.
 XX
 XX PA Whyte D, Manning G, Caenepeel S;
 XX
 XX DR WPI; 2004-122753/12.
 XX
 XX DR P-PSDB; ADJ96620.
 XX
 XX PT New nucleic acid molecule encoding a kinase polypeptide, useful for
 PT preparing a composition for treating diseases or disorders, e.g., cancer,
 PT or neurological, immunological or inflammatory disorders.
 XX

PS Example 1; SEQ ID NO 11; 366pp; English.

This invention relates to a novel isolated, enriched or purified nucleic acid molecule that encodes a kinase polypeptide. Specifically, it relates to human tyrosine and serine/threonine protein kinases (PTKs and STK's), CC as well as protein kinase-like enzymes. The present invention describes CC screening methods to identify agonists, antagonists and antibodies that CC can be used to modulate the activity or function of the mammalian kinase CC enzymes. As such, these compositions can be used for gene therapy CC purposes to treat diseases or disorders including cancer, immune-related CC diseases, cardiovascular disease, brain or neuronal associated disease, CC metabolic and inflammatory disorders. Accordingly, they exhibit CC cytosolic, neuroprotective, immunomodulator and antiinflammatory CC activities. This polynucleotide sequence is a human kinase DNA sequence CC of the invention.

SQ Sequence 3463 BP; 732 A; 1057 C; 919 G; 755 T; 0 U; 0 Other;

Query Match 71.7%; Score 1357; DB 12; Length 3463;

Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

OY	47	CTCCGGCCTGGCCTCCGGAGAGCGCCGGCCGCTGGCCGGA	CGGGCTCATCAAGTGCCTA	106
Db	223	CCATCCTCCCGCCGCAAGCTAGCCCGCCCGCTGGCGGA	AGGGCTGATCAAGTGCCTCA	282
OY	107	AACTCTGATGAAAGAGAGCGCGGTGAAGCGGCACATCA	CAAAACAACAATCGCGGCACC	166
Db	283	AGCCCTATGAAAGAAAGCAGGGCGGTGAAGCGGCAC	CCACCAAGACCAAACTGCGGCACC	342
OY	167	GCTACGAATTCCTGGAGACGCTGGGCAAAGGGCACTTA	CGGGAAGGTGAAGAAAGCAGAG	226
Db	343	GCTACGAATTCCTGGAGACCGCTGGGCAAAGGGCACTTA	CGGGAAGGTGAAGAAAGGCGCGGG	402
OY	227	AGAGCTGGGGGGGTGGGTGGCATCAAGTCATAGAAAGAA	CAAAATTCAAAGATGAGC	286
Db	403	AGAGCTGGGGGGGCTGGGTGGCATCAAGTCATCGAAGGA	CAAAATTCAAAGATGAGC	462
OY	287	AGAGTCTGCTGCACATACGAGAGGAGATTGAGATCATGT	CTTCACTCAACCAACCCGCACA	346
Db	463	AAGATCTGATGCAATACGAGAGGAGATTGAGATCATGT	CTCACTCAACCAACCTTCACA	522
OY	347	TCATTTGCCATCCATGAAGTGTTTGGAATAGCAGCAAG	ATTGTGATTCATGAGATAG	406
Db	523	TCATTTGCCATCCATGAAGTGTTTGGAATAGCAGCAAG	ATTGTGATTCATGAGATAG	582
OY	407	CCAGCCGAGGGGAGTCTGTATGATTAATCAATGAGAGCG	CCAGCGTGAGTGAGCGGAGACG	466
Db	583	CCAGCCGAGGGGAGCTTTATGATTAATCAATGAGAGCG	CCAGCGTGAGTGAGCGGAGACG	642
OY	467	CCAGGCAATTTCTTCCGACAGATCGGTCTGCGCTGCACA	CTTACTGCCACAGAAACGGGATCG	526
Db	643	CTAAGCAATTTCTTCCGAGAGATCGGTCTGCGCGTGACA	TTATTCGATCAGAAACAGATTTG	702
OY	527	TTCAACCGAGATCTCAAGCTGAAAAACATCTT	TAGATGCCAATGGAACATCAAGATTG	586
Db	703	TTCAACCGAGATCTCAAGCTGAAAAACATCTT	TAGATGCCAATGGAACATCAAGATTG	762
OY	587	CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTT	CTCTCCAGACGTTCTGTGGGA	646
Db	763	CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTT	CTCTCCAGACATTTCTGTGGGA	822
OY	647	GCCCTCTCTAGCCCTCGCTGAGATGATCAACGGGAAC	CCCTATGTGGAGCCACAGAGTGG	706
Db	823	GCCCCCTCTATGCTCTCGCCAGAGATTTGTCAATGGGA	AGCCCTTACACAGGCCACAGAGTGG	882
OY	707	ACAAGCTGATCTCTGGGCGTTCTCTGTACATCTGTG	ATCAGTGCACATGACCATGCTTTGACG	766
Db	883	ACAAGCTGATCTCTGGGCGTTCTCTGTACATCTGTG	ATCAGTGCACATGACCATGCTTTGATG	942
OY	767	GGCAGGATCATAAAACACTGGTGAAGCAATTCAGTA	ACGGGGCTTAAACCTGAGCCGCCCA	826
Db	943	GGCAATGACATTAAGTCTCTAGTGAACAAGATCAGAA	ACGGGGGCTTACCGGGAGGCCACTTA	1002

QY	827	AGCCCTCCGAGTCCCTGTGGCCCTGATATCCGGTGGCTGTTAATGGTGAACCCACCCGTCGGG	886
Db	1003	AACCCCTGATGCTGTGGCTGTGATCCGGTGGCTGTTAATGGTGAACCCACCCGTCGGG	1067
QY	887	CCACACTGAGAGATGTAGCCAGTCAATTGGTGGGTGAACGGGGATTACACACCGAGATGG	946
Db	1063	CCACCTTGAGAGATGTGGCCAGTCACTGGTGGGTGAACGGGGATTACCGCACCCGAGTGG	1122
QY	947	GGAAGACAGAAACCCCTGCGTGAAGGGGTGGACCTTAGTGTGACTTTTGGCCGGGCTTCCA	1000
Db	1123	GAGACAGAGAGGCTCCGATGAGGGTGGGACCCCTGGAGATGACTCTGGCCCGCTCTCA	1187
QY	1007	TGGGAGACMGGTTATGTCGTCTTCGGGCCCCCTCTGAGAAATGAGACCAAGGTGTGA	1066
Db	1183	TGGCTGACTGGCTCCGGGCTTCTTCCGCCCTCTCTGAGAAATGGGCGACAGTGTGA	1244
QY	1067	GCTTCTTCAAGCAGACGTCGCCGGAGGTGAAGCATTGACTTGGGCTGAGCCGACAC	1120
Db	1243	GCTTCTTCAAGCAGATGACACTGGTGGGGAGAACACCCCTGGCTGGAGCCGACG	1307
QY	1127	ATTCTTAAAGAAATGCCGAAAGAGATTTGACATGGCTCAAAATTTGACAGTTGACCCGG	1188
Db	1303	ATTGCTCAAGAAATGCCGAAAGAGATTTGACATGGCCAGTCTTTCACAGTGAACAGG	1367
QY	1187	CTGAGATTAACCTTTCTCGCCCTGGCAAGACGCTTTAAGCTTCCGAAAGGACTTCTCA	1246
Db	1363	CTGATGACATGTCCCATGCGCCCTGGCAAGACACTCAAGCTGCAAAAGGCAATTCTCA	1422
QY	1247	AGAAAGAAATGCTCTTACTCTGTTCAGGGGAGGTGACAGAGAGACCTCAGAACTCAGACGG	1306
Db	1423	AGAAAGAAATGTCAAGCTCTCTCAGAAAGGGGTGACAGAGAGACCTCTCGAGCTCAAGCCAA	1482
QY	1307	TGCCGTATTAATCCAGGGGAGCGTGGCCCTGCTGATCCCTGCTGCCCAAGAAAGGCAATCC	1366
Db	1483	TCCCTGCAAGCCCGAGGCGAGGCTGCCC-----CCCTGCTCCCAAGAGGCAATTC	1533
QY	1367	TTAAGAACTTCGACAGCGTGAATCTGGTTACTCTCTCTCCAGAGCCACGCGAGTCTG	1428
Db	1534	TCAAGAAACCCCGACAGCGGAGTCTGGCTACTACTCTCTCCGAGCCAGTGAATCTG	1592
QY	1427	GGGAATCTTTAGAGCGCAGTATGTGTTTGTGAATGGGAGACCCCGTGAAGCAAGTCTC	1488
Db	1594	GGAGACTTTTGAAGCAGGACGATGTTTGTGATGGGGAATCCAAAGACGAAAGCTTC	1653
QY	1487	CACAGGCTTCAAGGGCTCTCTCTCCACGCAAGGAGATTTCTAAACTCATGAGCAAGTCT	1546
Db	1654	CGAAGCTTCAAGGGGTGCTCTCTCATGCAAGAGATCTCTAAACTCATGAGCAAGTCT	1711
QY	1547	CCCGACAGCTTTGAGAGGCACTACCCCTAGACCTTTTGGCTCCCTGGAACCACTGGCT	1606
Db	1714	CCCAAGACGCTTTGAGAGCTCGCGGCCCCCAACACTTCGGCTCCCTGATGAATCGGCC	1772
QY	1607	CCTCCCAATCTCGACGCCCGGCGCCAGCGGCCCTCAGAGGGCGTGAAGTGAAGACAGATCC	1666
Db	1774	CACCTTCGCCCTTGCGCCCGGCGACGCGACCTCAGAGGGCTGTGAAGCGAGACAGATCC	1833
QY	1667	TGTCCTCGAGTCTTTTGAACCAATTTGACTTGCCTGAACGTCCTTCCGAAACCCCACTGA	1726
Db	1834	TGTCCTCGAGTCTTTTGAACCAAGCTTGAACCTTGAACCGCTCCCAAGCCCCCACTGC	1892
QY	1727	GGGCTGTGTGTCTGTGACAACTTGAGGGGGCTTTGAGCAAGCTTCCTCAAG-----	1788
Db	1894	GGGCTGTGTGTCTGTGACAACTTCAAGGGGCTTTGAGAGACCCCCCTCAAGGGGCCCTG	1953
QY	1781	-----GTCTGAACGATGTGTGGGAGGAATCTTTGGGGGATAGCTGTCTTCTCGACAG	1834
Db	1954	GAACTGTGTGATGGCGCTGGGGGAGATCTTTTGGGGGACAGCTGCTTTTCCCTGACAG	2013
QY	1835	ACTGCAAGAGATGTGACGCTTACAGACAGCCCTTAGAATGTGCTCAAGCTCAAGCTCAGC	1893
Db	2014	ACTGCAAGAGATGTGACGCACTTACGACAGGCACTGAAGGCTGTGCTCAAGCTCAAGCTCAGC	2072

RESULT 12
AB211333
ID AB211333 standard; cDNA; 2043 BP.
XX
AC AB211333;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 215.
XX
KW Human; genome mapping; gene therapy; food supplement; virtue; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antithetic; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002MO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RM, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drennec RT;
XX
DR WPI; 2002-759812/82.
XX
DR P-PSDB; ABP69116.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 215; 1012bp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB21119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP63849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from Wipo at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2043 BP; 418 A; 659 C; 592 G; 374 T; 0 U; 0 Other;

Query Match 71.6%; Score 1355.4; DB 6; Length 2043;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CTTCCGCGCTGCTCGGAGAGCGCCCGGCTGCGGAGCGGCTCATCAAGTCGGCTTA 106
DB 140 CCACTCCCTGCGCGAGAGCTAGCCCGCGCTGCGGAGAGGCGCTATCAAGTCGCCCA 199

QY 107 AACCTGATGAAGAGACAGCGCGGTGAAGCGGACCTCATCAAAACATAACCTGGGCAAC 166
DB 200 AGCCCTTAATGAAGAGACAGCGCGGTGAAGCGGACCTCATCAAAACATAACCTGGGCAAC 259
QY 167 GCTACGAGTTCCTGAGAGCGCTGGGCAAGGGCACTTAACGGAAAGGTGAAGAGGACGAG 226
DB 260 GCTACGAGTTCCTGAGAGCGCTGGGCAAGGGCACTTAACGGAAAGGTGAAGAGGCGGCG 319
QY 227 AGAGCTCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286
DB 320 AGAGCTCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 287 AGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 346
DB 380 AAGATCGATGACATACGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
QY 347 TCATTCGATCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 406
DB 440 TCATTCGATCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
QY 407 CCAAGCGAGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 466
DB 500 CCAAGCGAGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
QY 467 CCAAGCGATTTCTTCCGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
DB 560 CTAGGATTTCTTCCGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
QY 527 TTCAACGAGATCTCAAGCTGAGAAACATCTTCTAATGTCATGAAATCAATCAAGTTTG 586
DB 620 TCACCGAGATCTCAAGCTGAGAAACATCTTCTAATGTCATGAAATCAATCAAGTTTG 679
QY 587 CTGACTTTGGCTCTCTCAACCTGACCAAGAGGCAAGTTCTCTCAACAGCTTCTGTTGGA 646
DB 680 CTGACTTTGGCTCTCTCAACCTGACCAAGAGGCAAGTTCTCTCAACAGCTTCTGTTGGA 739
QY 647 GCCCTCTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
DB 740 GCCCTCTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
QY 707 AAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
DB 800 AAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 859
QY 767 GGCAGGATCAATAAACAAGCTGAGAGCAATCAAGTAAAGGAGCTTAACCTGAGCGGCGCA 826
DB 860 GGCAGGATCAATAAACAAGCTGAGAGCAATCAAGTAAAGGAGCTTAACCTGAGCGGCGCA 919
QY 827 AGCGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
DB 920 AACCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
QY 887 CCAACCTGAGAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 946
DB 980 CCAACCTGAGAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1039
QY 947 GGAACAAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1006
DB 1040 GGAACAAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
QY 1007 TGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
DB 1100 TGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1159
QY 1067 GCTTTCTTAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
DB 1160 GCTTTCTTAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
QY 1127 ATTCTCTTAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186
DB 1220 ATTCTCTTAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279
QY 1187 CTGAGGATTAACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246

DB 1280 CTATGACCTGCTCCATGCGCTGGCAAGCACTCAAGCTCCAAAGGCGCTTCTCA 1339
1247 AGAAAAATGCTCTACTCTGTCAGGGAGGTACAGAGAGCCCTCAGAACTCAGACGG 1306
DB 1340 AGAAGAAAGTGTGAGCTCTGCAAGAGGGGTACAGAGAGCCCTCGAGTCTAGCCCA 1399
1307 TGGCTGATCTCCAGGGGAGCTGTCCCTGCTGATCCCTGCTCCAGAGAAAGCATCC 1366
DB 1400 TCCCTGAGACCCAGGGGAGCTGCCCC-----CCCTGCTCCCAAGAGGCAATTC 1450
1367 TTAAGAGTCTGACAGCGTGAATCTGTTACTACTCTCTCCAGAGCCCAAGGAGCTG 1426
DB 1451 TCAAGAGCCCAAGAGCGAGTGTGCTACTACTCTCTCCAGAGCCCAAGGAGCTG 1510
1427 GGGAACTCTTGAAGCGCAGTGTGTTGTTGAGTGGGAGCCCGGAGCAGAGCTC 1486
DB 1511 GGGAGCTCTTGAAGCGAGGAGAGTGTGTTGAGTGGGAGTCCCAAGAGCAGAGCTC 1570
1487 CACAGGCTTCAAGGCTCTCTCTTCCAGCGCAAGGAGCTTCAAACTCAATGGCAAGTCT 1546
DB 1571 CGAAGCTTCAAGGCTCTCTCTTCCAGCGCAAGGAGCTTCAAACTCAATGGCAAGTCT 1630
1547 CCGGCAAGGCTTGAAGAGCACTAACCCCTGACCTTGGCTCCCTGAGCACTGAGCT 1606
DB 1631 CCGAGAGGCTTGAAGAGCTGAGCGGCGCCCACTTGGCTCCCTGAGCACTGAGCT 1690
1607 CTTCCATCTTGAAGCGGCGGAGCGGCGGCTTGAAGGAGTGTGAGAGCAGATCC 1666
DB 1691 CACTGCGGCGGCGGCGGCGGAGCGGAGCGGCTTGAAGGAGTGTGAGAGCAGATCC 1750
1667 TGTCTCTCAGAGCTCTTGAAGCAATTTGAGCTTGGCTTCCGAAACCCGACTGA 1726
DB 1751 TGTCTCTCAGAGCTCTTGAAGCAATTTGAGCTTGGCTTCCGAAACCCGACTGA 1810
1727 GGGGCTGTGTGTCTGAGCAACTGAGGGGCTTGAAGCGGCTTCCGAAAG----- 1780
DB 1811 GGGGCTGTGTGTCTGAGCAACTGAGGGGCTTGAAGCGGCTTCCGAAAG----- 1780
1781 -----GTCTGAAGCGATGTGAGCAAGATCTTGGGAGTAGCTGTTCTCTGACAG 1834
DB 1871 GAGCTGCGCTGAGGCGGTGGCGGAGAGATCTTGGGAGCAAGTGTCTTCTGACAG 1930
1835 ACTGCCAAGAGTGTGAGCTGACCTTACAGAGAGCGCTTGAAGTGTCTTCAAGCTCAGC 1893
DB 1931 ACTGCCAAGAGTGTGAGCAAGCTTACAGAGAGCGCTTGAAGTGTCTTCAAGCTCAGC 1989

RESULT 13
ADM43851
ID ADM43851 standard; cDNA; 2501 BP.
AC ADM43851;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #215.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
XX
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PR 05-MAR-2002; 2002WO-US005095.
XX
PR 20-AUG-2002; 2002US-00252551.
XX

PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
XX
PI Tang YT, Xue A, Drmanac RT,
XX
DR WPI; 2004-238579/22.
XX
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
PS Disclosure; SEQ ID NO 215; 51pp; English.
XX
XX The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
XX
SQ Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;
XX
Query Match 71.6%; Score 1355.4; DB 12; Length 2501;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;
XX
QY 47 CTTGCGGCTTGGAGAGCGCGGCGGCTGCGGAGCGGCTCATCAAGTCCCTA 106
DB 140 CCACTCCCTCGGCGGAGAGCTAGCCGCGCTGCGGAGGAGCTGATCAAGTCCGCCA 199
QY 107 AACCTGATGAAGAGAGAGCGGCTGAAGCGGAGCATCAACAACCACTCGGCGACC 166
DB 200 AGCCCTTAATGAAGAGAGAGCGGCTGAAGCGGAGCATCAACAACCACTCGGCGACC 259
167 GCTACAGATTCTGAGAGCGCTGGGCAAGGAGCACTTACGAGGAGTGAAGAGGAGCAGAG 226
DB 260 GCTACAGATTCTGAGAGCGCTGGGCAAGGAGCACTTACGAGGAGTGAAGAGGCGGAG 319
QY 227 AGAGCTCGGAGCGCTGTGGGCGATCAAGTCCATCAAGAGGAGCAAAATCAAGATGAGC 286
DB 320 AGAGCTCGGAGCGCGCTGTGGGCGATCAAGTCCATCAAGAGGAGCAAAATCAAGATGAGC 379
QY 287 AGAGTCTGTCGACATACGAGAGGAGATGAATCATGTCTTCACTCAACCAACCCCAACA 346
DB 380 AAGATCTGATGACATACGAGAGGAGATGAATCATGTCTTCACTCAACCAACCCCAACA 439
QY 347 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGATGATGATGATGATG 406
DB 440 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGATGATGATGATGATG 499
QY 407 CCAGCGAGGCGATCTGTATGATTACATGAGGCGGCGACGCTGAGTGAAGCGGAGAGC 466
DB 500 CCAGCGGAGGCGACCTTATGACTACATCAGGAGGCGGAGCGCTGAGTGAAGCGGAGAGC 559
QY 467 CCAGGATTTCTTCCGACAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526
DB 560 CTAGGCAATTTCTTCCGAGAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
QY 527 TTCACCGAGATCTCAAGCTGGAAGAAATCATCTTCTAGATGCAATGGAAGAAATCATGA 586
DB 620 TCACCGAGATCTCAAGCTGGAAGAAATCATCTTCTAGATGCAATGGAAGAAATCATGA 679
QY 587 CTGACTTTGAGCTCTTCAACCTGTACACAGAAAGCAAGTTCTTCAAGAGTTCTGTGGAG 646
DB 680 CTGACTTTGAGCTCTTCAACCTGTACACAGAAAGCAAGTTCTTCAAGAGTTCTGTGGAG 739
QY 647 GCGCTCTTCAAGCTGCTGCTGAGATGTCAAGGAGAGCCCTTATGTGGGCGGAGAGTGG 706
DB 740 GCGCTCTTCAAGCTGCTGCTGAGATGTCAAGGAGAGCCCTTATCAAGAGCCGAGAGTGG 799

QY 707 ACAGCTGCTCTGGGCGTCTCTCTGACATCTGTGATGAGCAACCATGCCCTTTGACG 766
DB 800 ACAGCTGCTCTGGGCGTCTCTCTGACATCTGTGATGAGCAACCATGCCCTTTGATG 859
QY 767 GGCAGGATCTATAAACACTGTGAGCAAAATCACTAACCGGGCTTTACCGAGCGGCCA 826
DB 860 GGCATGACATAAAGATCTTGTGTAACAGATCAAGCAACGGGGCTTACCGGAGCACTTA 919
QY 827 AGCGCTCGAATGCGCTGATGCTGATCCGGTGGCTGTTAATGAGTAAACCCACCGTGGG 886
DB 920 AACCTCTGATGCTGTGGCTGATCCGGTGGCTGTTAATGAGTAAACCCACCGCGGG 979
QY 887 CCACTGAGGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 980 CCACTGAGGAGATGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1039
QY 947 GGGAAACGAAAGCCCTGCTGAGGGGTGGGACCTGATGATGATGATGATGATGATGATG 1006
DB 1040 GAGAGCAGAGAGCTCCGATGAGGGTGGGACCTGATGATGATGATGATGATGATGATG 1099
QY 1007 TGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 1100 TGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159
QY 1067 GCTTCTTCAAGCAGCAGTGTGCGGAGGAGTGAAGCACTGATGATGATGATGATGATGATG 1126
DB 1160 GCTTCTTCAAGCAGCAGTGTGCGGAGGAGTGAAGCACTGATGATGATGATGATGATGATG 1219
QY 1127 ATTCTCTTAAGAGTCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1186
DB 1220 ATTCTCTTAAGAGTCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1279
QY 1187 CTGAGGATACCTCTCTCTGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
DB 1280 CTGAGGATACCTCTCTCTGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339
QY 1247 AGAAAGAGTCTTACCTGCTGAGGAGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
DB 1340 AGAAAGAGTCTTACCTGCTGAGGAGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1399
QY 1307 TGCCTGATCTTCCAGGAGGAGGAGTGTCTGCTGATCTCTGCTGCTGCTGCTGCTGCTG 1366
DB 1400 TGCCTGATCTTCCAGGAGGAGGAGTGTCTGCTGATCTCTGCTGCTGCTGCTGCTGCTG 1450
QY 1367 TTAAGAGTCTTCCAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1426
DB 1451 TCAAGAGGAGGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1510
QY 1427 GGGAACTCTTAAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1486
DB 1511 GGGAACTCTTAAGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1570
QY 1487 CAGAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1546
DB 1571 CAGAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1630
QY 1547 CCGGACAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1606
DB 1631 CCGGACAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1690
QY 1607 CCGGACAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1666
DB 1691 CCGGACAGGCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1750
QY 1667 TGTCTCTGAGTCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1726
DB 1751 TGTCTCTGAGTCTTGAAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1810
QY 1727 GGGGCTGTGTCTGTGAGCAACTGAGGGGGCTTTGAGAGAGCTTCTCAGAGAG----- 1780
DB 1811 GGGGCTGTGTCTGTGAGCAACTGAGGGGGCTTTGAGAGAGCTTCTCAGAGAGGGCCCTG 1870

QY 1781 -----GTCTGAAGCCATGCTGAGAGATCTTTGGGGGATAGTCTTTCTCTGACAG 1834
DB 1871 GAAGCTGCTGAGAGGCGCTGGCGGAGAGATCTTTGGGGGACAGCTGTCTTCCCTGACAG 1930
QY 1835 ACTGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1893
DB 1931 ACTGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1989

RESULT 14
ABX71420
ID ABX71420 standard; cDNA; 3443 BP.
XX
AC ABX71420;
XX
DT 14-Apr-2003 (first entry)
XX
DE Human cell cycle-associated cDNA from clone DKFZpht3e3_7j3.
XX
KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
OS Homo sapiens.
XX
PN W0200112659-A2.
XX
PD 22-FEB-2001.
XX
PE 18-AUG-2000; 2000MO-IB001496.
XX
PR 18-AUG-1999; 99US-014949P.
XX
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
XX
DR P-PSDB; ABUS3319.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Claim 1; Page 942-943; 1095BP; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence encodes a polypeptide
CC described in the disclosure of the invention
XX
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 71.6%; Score 1355.4; DB 5; Length 3443;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CTTCCGCGCTGGGCTGAGAGGCGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 106
DB 162 CCACTCTCTGAGGCGCGAGAGCTAGCCGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 221
QY 107 AACCTGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 166
DB 222 AGCCCTTAATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 281
QY 167 GCTACGAGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 226
DB 282 GCTACGAGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 341

QY 227 AGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGC 286
 DB 342 AGAGCTGGGGGCGCTGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGC 401
 QY 287 AGAGTCTGTGCATATCGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCA 346
 DB 402 AAGATCTGATGCAATACGAGGAGATGAGATCATGTCTCACTCAACCAACCCCA 461
 QY 347 TCATGTGCATTCATGAAAGTGTGTAAGATGACCAAGATTTGATGTCTGATGATG 406
 DB 462 TCATGTGCATTCATGAAAGTGTGTAAGATGACCAAGATTTGATGTCTGATGATG 521
 QY 407 CCAGCCGAGGCGATCTGTATGATTAATCAATCAAGTGGCGGCAACGCTGAGGCGGAG 466
 DB 522 CCAGCCGAGGCGATCTGTATGATTAATCAATCAAGTGGCGGCAACGCTGAGGCGGAG 581
 QY 467 CCAAGCATTTCTTCCGACAGATCTGTCTGCTGCACTAATGCAACAGACGGAATG 526
 DB 582 CTAGGCAATTTCTTCCGACAGATCTGTCTGCTGCACTAATGCAACAGAGATG 641
 QY 527 TTCAACGAGATTCATAGCTGGAACATCTTCTGATGATGCAATGAAATCAAGATG 586
 DB 642 TCACCCGAGATTCATAGCTGGAACATCTTCTGATGATGCAATGAAATCAAGATG 701
 QY 587 CTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGA 646
 DB 702 CTGACTTTGGCTCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGA 761
 QY 647 GGCCTCTCTAAGCCTCGCTGAGATAGTCAACGGAAGCCCTATGAGGCGCAGAGTGG 706
 DB 762 GGCCTCTCTAAGCCTCGCTGAGATAGTCAACGGAAGCCCTATGAGGCGCAGAGTGG 821
 QY 707 ACAGCTGTCTCTGAGGCGTTCTCTGTATCATCTGTGATGCAATGCAATGCAATG 766
 DB 822 ACAGCTGTCTCTGAGGCGTTCTCTGTATCATCTGTGATGCAATGCAATGCAATG 881
 QY 767 GGCAGATCATATAACCTGTGAGGAATCAAGTACGAGGCTTACGTAAGGCGGCCA 826
 DB 882 GGCAGATCATATAACCTGTGAGGAATCAAGTACGAGGCTTACGTAAGGCGGCCA 941
 QY 827 AGCGCTCGATGCTGTGGCTGATCCGCTGCTGTATATGATGAAACCCACCGCTCGG 886
 DB 942 AACCTCTGATGCTGTGGCTGATCCGCTGCTGTATATGATGAAACCCACCGCTCGG 1001
 QY 887 CCAACTGTGAGATGATGCAATGATGATGATGATGATGATGATGATGATGATG 946
 DB 1002 CCAACTGTGAGATGATGCAATGATGATGATGATGATGATGATGATGATGATG 1061
 QY 947 GGAAGACGAAAGCCGCTGAGAGGCTGAGCAACCTTATGATGATGATGATGATGATG 1006
 DB 1062 GGAAGACGAAAGCCGCTGAGAGGCTGAGCAACCTTATGATGATGATGATGATGATG 1121
 QY 1007 TGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
 DB 1122 TGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1181
 QY 1067 GCTTTCTTCAAGACGATGATGATGATGATGATGATGATGATGATGATGATG 1126
 DB 1182 GCTTTCTTCAAGACGATGATGATGATGATGATGATGATGATGATGATGATG 1241
 QY 1127 ATTCTCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
 DB 1242 ATTCTCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1301
 QY 1187 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
 DB 1302 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
 QY 1247 AGAAAAAGTCTTACTGCTGTCAAGGAGGATACAGAGAACCTCAGAGAACTCAGACCG 1306
 DB 1362 AGAAAAAGTCTTACTGCTGTCAAGGAGGATACAGAGAACCTCAGAGAACTCAGACCG 1421
 QY 1307 TGCTGATACTCCAGGAGGCTGTCTCTGTATCTCTGCTCCAGGAAAGGATCC 1366

DB 1422 TCCCTGGAGCCCAAGGACAGCTGCC-----CGCTGCTCCCAAGAGGCAATTC 1472
 QY 1367 TTAAGAAGTCTCGACAGCGTGAATCTGTACTACTCTTCTCCAGAGCCGAGGATCTG 1426
 DB 1473 TCAAGAAAGCCCGACAGCGGATGCTGGCTACTACTCTCTCCGAGCCCAAGATCTG 1532
 QY 1427 GGGATCTTTAAGCCGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1486
 DB 1533 GGGATCTTTAAGCCGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1592
 QY 1487 CACAGGCTTACAGGCTCTCTCTCCAGGAGGCAATTCGAACTCAATGGAAGTTCT 1546
 DB 1593 CGAGGCTTACAGGCTCTCTCTCCAGGAGGCAATTCGAACTCAATGGAAGTTCT 1652
 QY 1547 CCGCAGAGCTTTAAGAGGCACTACCTAGACATTTGGCTCTCCGACCACTGGCT 1606
 DB 1653 CCGCAGAGCTTTAAGAGGCACTACCTAGACATTTGGCTCTCCGACCACTGGCT 1712
 QY 1607 CTTCCATCTGCAAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1666
 DB 1713 CACTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1772
 QY 1667 TGTCTCCAGGCTCTTGAACCAATGGAATGGAATGGAATGGAATGGAATGGAATG 1726
 DB 1773 TGTCTCTGAGTCTTGTGACAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1832
 QY 1727 GGGGCTGT 1780
 DB 1833 GGGGCTGT 1892
 QY 1781 -----GTGTGAAGCAATGATGAGGAAATCTTGGGGGATGATGCTTTTCTGACAG 1834
 DB 1893 GAAGTGTCTGTGAGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1952
 QY 1835 ACTGCCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1893
 DB 1953 ACTGCCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2011
 RESULT 15
 ADF76964
 ID ADF76964 standard; cDNA; 3443 BP.
 XX
 AC ADF76964;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Novel human secreted and transmembrane protein cDNA Segid 639.
 XX
 KW human; PRO; membrane bound protein; membrane bound receptor;
 KW cell proliferation; cell migration; cell differentiation;
 KW mitogenic factor; survival factor; cytotoxic factor;
 KW differentiation factor; neuropeptide; hormone; cell receptor;
 KW receptor-ligand interaction; cytosolic; chondrocyte; tumour; ss; gene.
 XX
 OS Homo sapiens.
 XX
 PN M02003072035-A2.
 XX
 PD 04-SEP-2003.
 XX
 PF 21-FEB-2003; 2003MO-US005241.
 XX
 PR 22-FEB-2002; 2002US-0359461P.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR,
 PI Williams PM, Wood WI, Wu TD,
 XX
 DR WPI; 2003-721702/68.
 DR P-PSDB; ADF76965.

XX New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT diabetes mellitus.

PS Claim 2, SEQ ID NO 639, 918bp, English.

XX This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neuropeptides and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytosolic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is a cDNA sequence which encodes a human PRO protein of the
CC invention.

SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 71.6%; Score 1355.4; DB 10; Length 3443;

Best Local Similarity 84.0%; Pred. No. 0; Mismatches 276; Indels 21; Gaps 2;

Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 47 CCTCCGCGCTGCTCGGAGAGCGCGCGCTGCGGAGCGGCTCATCAAGTCGCTA 106
DB 162 CCACTCCCTCGGCGGAGAGCTACCGCGCGCTGCGGAGCGGCTCATCAAGTCGCTA 221
QY 107 AACCTGATGAGAGAGAGCGGCTGAGCGGCACTCAACAACCAACTGCGGCAAC 166
DB 222 AGCCCTTAATGAAGAGAGAGCGGCTGAGCGGCACTCAACAACCAACTGCGGCAAC 281
QY 167 GCTACGAGTTCCTGAGAGAGCGGCTGAGCGGCACTCAACAACCAACTGCGGCAAC 226
DB 282 GCTACGAGTTCCTGAGAGAGCGGCTGAGCGGCACTCAACAACCAACTGCGGCAAC 341
QY 227 AGAGCTGCGGCGCTGCTGCTGCTCATCAAGTCATCAAGAGAGAGAGAGAGAGAG 286
DB 342 AGAGCTGCGGCGCTGCTGCTGCTCATCAAGTCATCAAGAGAGAGAGAGAGAGAG 401
QY 287 AGAGCTGCTGCTGCTGCTGCTGCTCATCAAGTCATCAAGAGAGAGAGAGAGAGAG 346
DB 402 AAGATCTGATGCAATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
QY 347 TCATTCGCTCATGAG 406
DB 462 TCATTCGCTCATGAG 521
QY 407 CCAAGCCGAGGCGATCTGATGATTAATCAATCAATGAGCGGCGAGAGAGAGAGAGAGAG 466
DB 522 CCAAGCCGAGGCGATCTGATGATTAATCAATCAATGAGCGGCGAGAGAGAGAGAGAGAG 581
QY 467 CCAAGCCGAGGCGATCTGATGATTAATCAATCAATGAGCGGCGAGAGAGAGAGAGAGAG 526
DB 582 CTAAGGCAATTTCTTCGAG 641
QY 527 TTCAACGAGATCTCAAG 586
DB 642 TCACCGAGATCTCAAG 701
QY 587 CTGACTTTGAGCTCTCAACCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646

DB 702 CTGACTTTGAGCTCTCAACCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761
QY 647 GCCCTCTCAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 706
DB 762 GCCCTCTCAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 821
QY 707 ACAAGTGTCTCTGAGGAGGCTTCTCTGATCAATCTGATGATGATGATGATGATGATGATG 766
DB 822 ACAAGTGTCTCTGAGGAGGCTTCTCTGATCAATCTGATGATGATGATGATGATGATGATG 881
QY 767 GGCAGAGATCTAAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 826
DB 882 GGCAGATCTAAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 941
QY 827 AGCGGTCCAGATGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 886
DB 942 AACCTCTGATGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1001
QY 887 CCAACTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
DB 1002 CCAACTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1061
QY 947 GGAAGACAGAGAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1006
DB 1062 GAGAGCAGAGAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1121
QY 1007 TGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
DB 1122 TGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1181
QY 1067 GCTTCTTCAAGCAGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1126
DB 1182 GCTTCTTCAAGCAGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1241
QY 1127 ATTCTCTTAAGAGTCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1186
DB 1242 ATTCTCTTAAGAGTCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1301
QY 1187 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
DB 1302 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
QY 1247 AGAAGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
DB 1362 AGAAGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1421
QY 1307 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1366
DB 1422 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1481
QY 1367 TTAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1426
DB 1473 TTAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1532
QY 1427 GGAAGTCTTAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
DB 1533 GGAAGTCTTAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1592
QY 1487 CAGAGGCTTCAAGGCTCTCTCTCAACCGAGAGGCAATTCATCAATGATGATGATGATG 1546
DB 1593 CAGAGGCTTCAAGGCTCTCTCTCAACCGAGAGGCAATTCATCAATGATGATGATGATG 1652
QY 1547 CCGGACAGGCTTCAAGGCTCTCTCTCAACCGAGAGGCAATTCATCAATGATGATGATG 1606
DB 1653 CCGGACAGGCTTCAAGGCTCTCTCTCAACCGAGAGGCAATTCATCAATGATGATGATG 1712
QY 1607 CTTCCCATCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1666
DB 1713 CACTTGCGCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1772
QY 1667 TGTCTTCAGAGTCTCTTTCATCAATGATGATGATGATGATGATGATGATGATGATGATG 1726

Db 1773 TGTCTCTGAGTCTTGTGACAGCTGGACTTGCCTGAACGGCTCCAGAGCCCCCACTGC 1832
QY 1727 GGGGCTGTGTGTCTGTGGAACAACCTGAGGGGCTTGAAGCAAGCTTCCCTCAGAAAG----- 1780
Db 1833 GGGGCTGTGTGTGTGTGGAACAACCTCAGGGGCTTGAAGAGCCCCCTCAGAGGGCCCTG 1892
QY 1781 -----GTCTGAAGCGATGGTGGCAGGAATCCTTGGGGGGAATAGCTGCTTTTCTGTGACAG 1834
Db 1893 GAAAGCTGCTGAGGCGCTGGGGCAGGAATCCTTGGGGGACAGCTGCTTTTCTGTGACAG 1952
QY 1835 ACTGCCAAGAGGTGACTGTGACAGCCTTACAGACAAGCCCTAGGAATCTGCTCAAGGCTCAGC 1893
Db 1953 ACTGCCAAGAGGTGACAGCCTTACCGACAGGCACTGAGGGTCTGCTCAAGGCTCAGC 2011

Search completed: May 15, 2005, 03:21:46
Job time : 989.386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 06:02:36 / Search time 819.38 Seconds
(without alignments)
11298.520 Million cell updates/sec

Title: US-09-980-464-4_COPY_123_2015

Perfect score: 1893
Sequence: 1 atgagatcggtgagccttact.....gaatctgtctcaagctcagc 1893

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sce:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	6	AR279568 Sequence
2	1893	100.0	2902	6	AR562226 Sequence
3	1888.2	99.7	2027	6	AX381046 Sequence
4	1872.2	98.9	2917	10	BC033302 Mus muscu
5	1859	98.2	2938	10	BC046833 Mus muscu
6	1686.6	89.1	2900	10	BC081899 Rattus no
7	1683.4	88.9	2026	6	AX380960 Sequence
8	1683.4	88.9	2929	6	AX380960 Sequence
9	1528.4	80.7	291762	2	AC125887 Rattus no
10	1360.2	71.9	1884	6	AX407118 Sequence
11	1360.2	71.9	3353	6	AX407118 Sequence
12	1357.6	71.7	3360	6	AX642961 Sequence
13	1357	71.7	1887	6	CQ714327 Sequence
14	1357	71.7	3395	6	CQ782778 Sequence
15	1357	71.7	3395	6	BD127309 Primer fo
16	1357	71.7	3395	9	AK074830 Homo sapi
17	1357	71.7	3402	9	BC017306 Homo sapi
18	1355.4	71.6	3443	6	AX086949 Sequence
19	1355.4	71.6	3443	6	AX780351 Sequence

20	1338.8	70.7	2291	6	AX056394 Sequence
21	1230	65.0	1833	6	AX803413 Sequence
22	1071.4	56.6	1186	6	AX380972 Sequence
23	943.4	49.8	1186	6	AX381029 Sequence
24	923.4	48.8	252978	2	AC126983 Rattus no
25	840.2	44.4	2616	6	AX430359 Sequence
26	840.2	44.4	2619	6	AX430468 Sequence
27	723	38.2	179414	9	AC098933 Homo sapi
28	719.8	38.0	6389	9	BM3806181 Homo sapi
29	651.8	34.4	1274	5	BM350299 Gallus ga
30	648	34.2	707	6	AX381050 Sequence
31	648	34.2	712	6	AX380990 Sequence
32	644.2	34.0	661	6	AX381052 Sequence
33	613.6	32.4	741	6	AX380991 Sequence
34	592	31.3	734	6	AX381051 Sequence
35	583.2	30.8	588	6	AX381056 Sequence
36	557.4	29.4	660	6	AX380992 Sequence
37	533	28.2	578	6	AX381054 Sequence
38	515	27.2	548	6	AX381055 Sequence
39	510.6	27.0	579	6	AX380994 Sequence
40	506.4	26.8	1866	6	AX213158 Sequence
41	506.4	26.8	2884	6	AX213160 Sequence
42	506.4	26.8	6825	6	CQ715220 Sequence
43	506.4	26.8	6828	6	CQ716374 Sequence
44	506.4	26.8	6828	6	AX213156 Sequence
45	506.4	26.8	6828	9	AB011109 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR279568 2902 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 4 from patent US 6514719.
ACCESSION AR279568
VERSION AR279568.1 GI:29714427
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2902)
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.
TITLES Methods for identifying compounds that alter kinase activity
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;
FEATURES
Source 1..2902
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1893; DB 6; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGATCGGTGAGCCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCCGCGCTGAGC 60
DB 123 ATGAGATCGGTGAGCCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCCGCGCTGAGC 182
QY 61 TCGGAGAGCGCCGCGCGCTGCGGAGCGGCGCTTCATCAAGTCGCTTAACTCTGATGAAG 120
DB 183 TCGGAGAGCGCCGCGCGCTGCGGAGCGGCGCTTCATCAAGTCGCTTAACTCTGATGAAG 242
QY 121 AAGCAGCGCGTGAAGGCGACCATCAACAACCACTCGGACCGCTACAGATTCCTG 180
DB 243 AAGCAGCGCGTGAAGGCGACCATCAACAACCACTCGGACCGCTACAGATTCCTG 302
QY 181 GAGACCGCTGGGCAAGGCGACCTACGAGGAGTGAAGAGGCAAGAGAGCTGGGGCGT 240
DB 303 GAGACCGCTGGGCAAGGCGACCTACGAGGAGTGAAGAGGCAAGAGAGCTGGGGCGT 362
QY 241 CTGGTGGCATCAAGTCCATCAGGAAGAACAATCAAAAGATGAGAGATCTGCTGCAC 300

Db 363 CTGGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAAATGAGCAGATCTGCTGCAC 422
 QY 301 ATACGAGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 360
 Db 423 ATACGAGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCAT 482
 QY 361 GAAGTGTGAGAAATGACGACGAAATTTGATGATGAGATGATGACGACGACGAT 420
 Db 483 GAAGTGTGAGAAATGACGACGAAATTTGATGATGAGATGATGACGACGACGAT 542
 QY 421 CTGTATGATTAATCATGATGAGCGGACGACGCTGATGAGCGGACGACGATTTCTTC 480
 Db 543 CTGTATGATTAATCATGATGAGCGGACGACGCTGATGAGCGGACGACGATTTCTTC 602
 QY 481 CGACGATGTGTCTTCCTGCACTA CTGCAACGAGACGAGATCTTCAACGAGATCTTC 540
 Db 603 CGACGATGTGTCTTCCTGCACTA CTGCAACGAGACGAGATCTTCAACGAGATCTTC 662
 QY 541 AAGCTGGAAGAACTCTTCTAGATGCAATGAGAAATCAAGATGCTGATCTTGGCTTC 600
 Db 663 AAGCTGGAAGAACTCTTCTAGATGCAATGAGAAATCAAGATGCTGATCTTGGCTTC 722
 QY 601 TCCAACTGTACCAAAAGGCAAGTCTCTCAGACGCTTCTGTGGAGCCTCTCTACGCC 660
 Db 723 TCCAACTGTACCAAAAGGCAAGTCTCTCAGACGCTTCTGTGGAGCCTCTCTACGCC 782
 QY 661 TGGCTGAGATATGATCAACGAGAAAGCCTTATGTGGGCCAGAGGTGACAGCTGTCTCTG 720
 Db 783 TGGCTGAGATATGATCAACGAGAAAGCCTTATGTGGGCCAGAGGTGACAGCTGTCTCTG 842
 QY 721 GGCCTTCTCTGTATCATCTGATGATGAGCAACATCCCTTTGACGGGAGAGATCATATA 780
 Db 843 GGCCTTCTCTGTATCATCTGATGATGAGCAACATCCCTTTGACGGGAGAGATCATATA 902
 QY 781 ACACTGTGTAAGCAATCATGTAACGAGGCTTACCTGAGCGGCCCAAGCGTCCGATGAC 840
 Db 903 ACACTGTGTAAGCAATCATGTAACGAGGCTTACCTGAGCGGCCCAAGCGTCCGATGAC 962
 QY 841 TGTGGCTGTATCCGCTGCTGTTAATGTGTAACCCCAACCCCTGCGGCACTGAGAGAT 900
 Db 963 TGTGGCTGTATCCGCTGCTGTTAATGTGTAACCCCAACCCCTGCGGCACTGAGAGAT 1022
 QY 901 GTAGCAGATCATTTGGGGGTCAACCTGGGGTTTACACCAACCGAGATCGGGGAAACGAAAGC 960
 Db 1023 GTAGCAGATCATTTGGGGGTCAACCTGGGGTTTACACCAACCGAGATCGGGGAAACGAAAGC 1082
 QY 961 CTGCGTGAAGGTGAGGACCTTAAGTGTGACTTTGCGCGGCTCCATGCGGACGTGTTA 1020
 Db 1083 CTGCGTGAAGGTGAGGACCTTAAGTGTGACTTTGCGCGGCTCCATGCGGACGTGTTA 1142
 QY 1021 CGTGGCTCTCGCGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTTCTCAAGCAG 1080
 Db 1143 CGTGGCTCTCGCGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTTCTCAAGCAG 1202
 QY 1081 CAGCGCGGAGGTGAGGACCTGTAACCTGGGCTGGAACCGCAATCTCTTAAGAG 1140
 Db 1203 CAGCGCGGAGGTGAGGACCTGTAACCTGGGCTGGAACCGCAATCTCTTAAGAG 1262
 QY 1141 TCCGAAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCCGCTGAGGATACCTCT 1200
 Db 1263 TCCGAAAGAGAAATGACATGCTCAAAATCTGCAAGTGAACCCGCTGAGGATACCTCT 1322
 QY 1201 TCTGGCCCTGCAAGACGACCTTAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1260
 Db 1323 TCTGGCCCTGCAAGACGACCTTAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1382
 QY 1261 ACCTGTGAGGGGAGGATACAGAGACCTTCAAGAACTAGACCGGTGCTGATCTCA 1320
 Db 1383 ACCTGTGAGGGGAGGATACAGAGACCTTCAAGAACTAGACCGGTGCTGATCTCA 1442
 QY 1321 GGGCAGCCTGTCTCTGCTGTATCTCCGCTCCCAAGAGAAAGCATCTTAAGAAAGTCTCA 1380
 Db 1443 GGGCAGCCTGTCTCTGCTGTATCTCCGCTCCCAAGAGAAAGCATCTTAAGAAAGTCTCA 1502

QY 1381 CAGCGTAATCTGTGTACTACTCTCTCAGAGCCCAAGGAGTCTGGGAACTTTAGAC 1440
 Db 1503 CAGCGTAATCTGTGTACTACTCTCTCAGAGCCCAAGGAGTCTGGGAACTTTAGAC 1562
 QY 1441 GCGAGTATGTGTTTGTGATGAGGAGCCCGGTGAGACAAAGTCTCAAGCTTTAGAG 1500
 Db 1563 GCGAGTATGTGTTTGTGATGAGGAGCCCGGTGAGACAAAGTCTCAAGGCTTTAGAG 1622
 QY 1501 CTCCTCTCCACCGCAAGGAGCAATCTCAAACTCAATGAGCAAGTCTCCGACAGCTTA 1560
 Db 1623 CTCCTCTCCACCGCAAGGAGCAATCTCAAACTCAATGAGCAAGTCTCCGACAGCTTA 1682
 QY 1561 GAAAGCACTACCTCTAGACCTTTGGCTCTCTGAGCAACCTGAGCTCTCCATCTGCA 1620
 Db 1683 GAAAGCACTACCTCTAGACCTTTGGCTCTCTGAGCAACCTGAGCTCTCCATCTGCA 1742
 QY 1621 GCCCGGCCAGCCGCCCTCAGAGGCTGTGATGAGGAGACAGATCTGTCTCCAGTCC 1680
 Db 1743 GCCCGGCCAGCCGCCCTCAGAGGCTGTGATGAGGAGACAGATCTGTCTCCAGTCC 1802
 QY 1681 TTTGACCAATTTGAACTTGTGCTGAACGCTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1740
 Db 1803 TTTGACCAATTTGAACTTGTGCTGAACGCTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1862
 QY 1741 GTGACCACTGAGGGGCTTTAGACAGCTCTCTCAAGAGTCTGAAGCGATGTGGAG 1800
 Db 1863 GTGACCACTGAGGGGCTTTAGACAGCTCTCTCAAGAGTCTGAAGCGATGTGGAG 1922
 QY 1801 GAATCCTTGGGGGATGATGCTGCTTTCTGACAGACTGCAAGAGTGACTGACAGCTAC 1860
 Db 1923 GAATCCTTGGGGGATGATGCTGCTTTCTGACAGACTGCAAGAGTGACTGACAGCTAC 1982
 QY 1861 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
 Db 1983 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 2015

RESULT 2
 AR562226 2902 bp DNA linear PAT 08-OCT-2004
 LOCUS AR562226
 DEFINITION Sequence 4 from patent US 6759223.
 ACCESSION AR562226
 VERSION AR562226.1 GI:53976143
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2902)
 AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.
 TITLE Calcium/calmodulin-dependent kinase
 JOURNAL Patent: US 6759223-A 4 06-JUL-2004;
 FEATURES location/Qualifiers
 source 1..2902
 /organism="Unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 1893; DB 6; Length 2902;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTGGTGGCTTACTCCAGGCGCCGAGCGAGCTCCCTGGGCTCCGCGCTGAGC 60
 Db 123 ATGAGTGGTGGCTTACTCCAGGCGCCGAGCGAGCTCCCTGGGCTCCGCGCTGAGC 182
 QY 61 TCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTGCCTTAAACTTGTGATGAG 120
 Db 183 TCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTGCCTTAAACTTGTGATGAG 242
 QY 121 AAGCAGGGGTGAGAGCGGACCATCAAAACAAACCTGCGGACCGCTAGAGATTCTG 180
 Db 243 AAGCAGGGGTGAGAGCGGACCATCAAAACAAACCTGCGGACCGCTAGAGATTCTG 302

181 GAGACGCTGGGAGAGGCACTTAACGGAAAGTGAAGAGCAAGAGAGCTCGGGCGT 240
DB GAGACGCTGGGAGAGGCACTTAACGGAAAGTGAAGAGCAAGAGAGCTCGGGCGT 362
QY 241 CTGTGGGCAATCAATGCTCATGAGAAAGCAAAATCAAAATGATGAGAGATCTGTGAC 300
DB 363 CTGTGGGCAATCAATGCTCATGAGAAAGCAAAATCAAAATGATGAGAGATCTGTGAC 422
QY 301 ATACGAGGAGAGTGAATGATCATGCTTCACTCAACACCCCAATCAATGATCATCAT 360
DB 423 ATACGAGGAGAGTGAATGATCATGCTTCACTCAACACCCCAATCAATGATCATCAT 482
QY 361 GAAGTGTGAGAAATGAGCAAGATTTGATGATGATGATGATGATGATGATGATGAT 420
DB 483 GAAGTGTGAGAAATGAGCAAGATTTGATGATGATGATGATGATGATGATGATGAT 542
QY 421 CTGTATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 543 CTGTATGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 481 CGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 603 CGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
QY 541 AAGCTGGAAGAAATCTTCTAGATGCAATGAAATCATCAAGATGCTGATCTTGGCTTC 600
DB 663 AAGCTGGAAGAAATCTTCTAGATGCAATGAAATCATCAAGATGCTGATCTTGGCTTC 722
QY 601 TCCAACTGTATCAAAAGGCAAGTCTTCCAGAGCTTGTGAGAGCTCTCTAAGCC 660
DB 723 TCCAACTGTATCAAAAGGCAAGTCTTCCAGAGCTTGTGAGAGCTCTCTAAGCC 782
QY 661 TGGCTGATATGATCAAGGAGAGCTTATGATGATGATGATGATGATGATGATGATGAT 720
DB 783 TGGCTGATATGATCAAGGAGAGCTTATGATGATGATGATGATGATGATGATGATGAT 842
QY 721 GGGCTTCTCTGTATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 843 GGGCTTCTCTGTATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 902
QY 781 ACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 903 ACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
QY 841 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 963 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
QY 901 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1023 GTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
QY 961 CTGCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1083 CTGCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
QY 1021 CGTGTCTCTCGGCGCCCTCTCTGAGATGAGCAAGTGTGAGCTTCTTCAAGCAG 1080
DB 1143 CGTGTCTCTCGGCGCCCTCTCTGAGATGAGCAAGTGTGAGCTTCTTCAAGCAG 1202
QY 1081 CAGGTGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1203 CAGGTGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
QY 1141 TCCGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1263 TCCGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
QY 1201 TCTGCGCTGAGCAAGAGAGCTTGAAGCTTCAAGAGAGCTTCTCAAGAGAGAGCTTCT 1260
DB 1323 TCTGCGCTGAGCAAGAGAGCTTGAAGCTTCAAGAGAGCTTCTCAAGAGAGAGCTTCT 1382

QY 1261 ACTGCTGAGGAGAGTATAGAGAGAGCTTCAAGAGAGCTTCAAGAGAGCTTCAAGAGAG 1320
DB 1383 ACTGCTGAGGAGAGTATAGAGAGAGCTTCAAGAGAGCTTCAAGAGAGCTTCAAGAGAG 1442
QY 1321 GGGCAGGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1443 GGGCAGGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
QY 1381 CAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1503 CAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
QY 1441 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1563 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
QY 1501 CTCTCTCTCAACCGCAAGGAGATTTCAAACTCAATGAGAGATTTCCCGCAAGCTTCA 1560
DB 1623 CTCTCTCTCAACCGCAAGGAGATTTCAAACTCAATGAGAGATTTCCCGCAAGCTTCA 1682
QY 1561 GAAGGCACTACCCCTAGACACTTGTGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1683 GAAGGCACTACCCCTAGACACTTGTGATGATGATGATGATGATGATGATGATGATGAT 1742
QY 1621 GCCCGGCGCAGCGCCCTCAAGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1743 GCCCGGCGCAGCGCCCTCAAGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1802
QY 1681 TTTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1803 TTTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
QY 1741 GTGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1863 GTGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
QY 1801 GAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1923 GAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
QY 1861 AGACAAAGCCCTTAGAATCTGCTCAAAAGCTCAGC 1893
DB 1983 AGACAAAGCCCTTAGAATCTGCTCAAAAGCTCAGC 2015

RESULT 3
AX381046 2027 bp DNA linear PAT 18-MAR-2002
LOCUS AX381046
DEFINITION Sequence 87 from Patent WO0212456.
ACCESSION AX381046
VERSION AX381046.1 GI:19575868
KEYWORDS
SOURCE
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.
Ampk-related serine/threonine kinase, designated smark
Patent: WO 0212456-A 87 14-FEB-2002;
1149336 ONTARIO INC. (CA)
FEATURES
source 1..2027
/organism="Mus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10095"
ORIGIN
Query Match 99.7%; Score 1888.2; DB 6; Length 2027;
Best Local Similarity 99.8%; Pred. 0;
Matches 1890; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGAGCTGGTGGCTTAATCTCAAGCGCCGAGCAGAGCTTCTCGGCTTCCGCTTGGCC 60

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.B., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, A.C., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22389257

JOURNAL

PUBMED

12477932

2 (bases 1 to 2917)

Strausberg, R.

Direct Submission

Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulgesed, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>

Series: IRAX Plate: 40 Row: d Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229181.

FEATURES

source

1..2917

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="MGC:30598 IMAGE:3663595"

/tissue_type="Mammary tumor metastasized to lung."

MMTV-LTR/Mnt1 Expression driven by an MMTV-LTR enhancer."

/clone_1ib="MCI CGAP_Lu30"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1..2917

/gene="1200013B22Rik"

/note="synonym: Snark"

/db_xref="locusID:74137"

/db_xref="MGI:1921387"

81..1976

/gene="1200013B22Rik"

/codon_start=1

/product="1200013B22Rik protein"

/protein_id="AAH33302.1"

/db_xref="MGI:233721086"

/db_xref="locusID:74137"

/db_xref="MGI:1921387"

CDS

KRHNNHNLHRYEPLTIGKTYGVKKARSSGRLVAIKSKIKDQDLAIR

REIEMSLNHPHIIAHEVPENSKIYVMEASRGDLYDISEPRLESDARHFE
ROIVSLAHKCHONGVIVHRDKEIILLDNGNITKADFGISNLYHNGKLOTCGSL
YASPELVNKPYPNGVDSGLVLLIYVHNGPDGDKHLVQIISGAAYRPP
PSDAGLILMLVNPTRATLEVDVSHMMVWGYTGVGEALREGGSDSGA
SNADMRSSRPLENGARVSPKOHVGGSTVGLRSHLSKSRKENDANOIO
GPDABTSSRPPKSSILKPKGILKXSSSTSSGSEVDPDLPVPPTGPVAVSL
PKGILKXSRKSESGTSSPESSEGLDASVFTGSPVPEAKESQASGLLRKGI
LKNGFSTALEGTTSTFGSLDQASSHPAPRPSFGAVERDSILSESSDQDL
PSRLPPTPLRGCVSYDNLRLGLOPSPESGKRWQESLSDGCSFLTDCOVYTAQAL
GICSKLS"

249..1001

/gene="1200013B22Rik"

/note="S_TKc Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"

/db_xref="CDI:cd00180"

ORIGIN

Query Match 98.9%; Score 1872.2; DB 10; Length 2917;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1880; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

1 ATGGAATCGGTGGCTTACTCCAGGCGCGAGCAGGCTCCCGGCTCCGCGCTGGCC 60
|||||
81 ATGGAATCGGTGGCTTACTCCAGGCGCGAGCAGGCTCCCGGCTCCGCGCTGGCC 140
|||||
61 TCGGAAGAGCGCCGCGCGCTGCGCGAGCGGCTCATGAAGTGGCTAACTCTGATGAAG 120
|||||
141 TCGGAAGAGCGCCGCGCGCTGCGCGAGCGGCTCATGAAGTGGCTAACTCTGATGAAG 200
|||||
121 AAGCAGCGCGTGAAGCGCGCATCATCAAAACACACCTCGGCGACCGCTACGAGTTCTG 180
|||||
201 AAGCAGCGCGTGAAGCGCGCATCATCAAAACACACCTCGGCGACCGCTACGAGTTCTG 260
|||||
181 GAGAGCGTGGCGAAGGCGACCTACGAGGAGTGAAGAGCGAGAGAGCTGGGGCGT 240
|||||
261 GAGAGCGTGGCGAAGGCGACCTACGAGGAGTGAAGAGCGAGAGAGCTGGGGCGT 320
|||||
241 CTGGTGCCATCAAGTCCATCAGGAAGAACAATATCAAGATGAGAGAGTCTGCTGAC 300
|||||
321 CTGGTGCCATCAAGTCCATCAGGAAGAACAATATCAAGATGAGAGAGTCTGCTGAC 380
|||||
301 ATACGAGGAGAGTGAATCATGTCTTCACTCAACCAACCCCAATCATTCGCATTCAT 360
|||||
381 ATACGAGGAGAGTGAATCATGTCTTCACTCAACCAACCCCAATCATTCGCATTCAT 440
|||||
361 GAAGTGTGAGAAATGACAGCAAGATTTGATGATGAGAGTATCCAGCGAGAGGAT 420
|||||
441 GAAGTGTGAGAAATGACAGCAAGATTTGATGATGAGAGTATCCAGCGAGAGGAT 500
|||||
421 CTGTATGATTAATCATGATGAGCGGCGACGCTGATGAGAGGAGCGAGAGTCTTCTTC 480
|||||
501 CTGTATGATTAATCATGATGAGCGGCGACGCTGATGAGAGGAGCGAGAGTCTTCTTC 560
|||||
481 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGAGCGAGTCTTCAACGAGATCTC 540
|||||
561 CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGAGCGAGTCTTCAACGAGATCTC 620
|||||
541 AAGCTGAAAACATCTCTTGAATGCGCATGAAAACATCAAGATTTGCTGACTTGGCTTC 600
|||||
621 AAGCTGAAAACATCTCTTGAATGCGCATGAAAACATCAAGATTTGCTGACTTGGCTTC 680
|||||
601 TCGACCTGTAACCAAGGCAAGTCTTCAAGAGTCTGATGAGAGCGCTCTCTAGCGCC 660
|||||
681 TCGACCTGTAACCAAGGCAAGTCTTCAAGAGTCTGATGAGAGCGCTCTCTAGCGCC 740
|||||
661 TCGCCTGAGATGATCAAGGGAAGCCCTATGTTGGGCGCAGAGGTGACAGAGTGTCTCTG 720
|||||

Db 741 TCGCCTGAGATATGCAACGGGAAGCCCTATGTGGGCCAGAGGTGACAGCTGTCTCTG 800
 Qy 721 GGGGTTCTCCTGTAACATCTCTGTGACATGAGCAATGCCCCCTTTTGAACGGGACGATCATATA 780
 Db 801 GGGGTTCTCCTGTAACATCTCTGTGACATGAGCAATGCCCCCTTTTGAACGGGACGATCATATA 860
 Qy 781 ACACTGTGAGCAATCAAGTAACGGGGCTTACCGTAGCGCCCAAGCGGTCCGATGCC 840
 Db 861 ACACTGTGAGCAATCAAGTAACGGGGCTTACCGTAGCGCCCAAGCGGTCCGATGCC 920
 Qy 841 TGTGACCTGATCCGGTGTCTGTTAATGTGAAACCCCAACCCGTGCGGCACACTGAGAGAT 900
 Db 921 TGTGACCTGATCCGGTGTCTGTTAATGTGAAACCCCAACCCGTGCGGCACACTGAGAGAT 980
 Qy 901 GTAGCCAGATCTATTGTGTGGCTCACTGGGCTTAAACCAACCGGAGTGGGGAACAGAAACC 960
 Db 981 GTAGCCAGATCTATTGTGTGGCTCACTGGGCTTAAACCAACCGGAGTGGGGAACAGAAACC 1040
 Qy 961 CTGCGTGAAGGTGGGACCCCTAGTGTGACTTTGGCCGGGCTCCANTGGGGACTGGTTA 1020
 Db 1041 CTGCGTGAAGGTGGGACCCCTAGTGTGACTTTGGCCGGGCTCCANTGGGGACTGGTTA 1100
 Qy 1021 CGTGGCTCTCGGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTCTTCAAGCAG 1080
 Db 1101 CGTGGCTCTCGGCCCTCTCTGAGAAATGAGCCAGGTGTGACGCTTCTTCAAGCAG 1160
 Qy 1081 CAGTCGCGGGAGGTGGAAGCACTGTACCTGGGCTGAGCGGCAACTTCTTTAAGAAAG 1140
 Db 1161 CAGTCGCGGGAGGTGGAAGCACTGTACCTGGGCTGAGCGGCAACTTCTTTAAGAAAG 1220
 Qy 1141 TCCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCCGGGTGAGATCTCTC 1200
 Db 1221 TCCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGTGACCCGGGTGAGATCTCTC 1280
 Qy 1201 TCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTCTCAAGAAAGTCTCTC 1260
 Db 1281 TCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTCTCAAGAAAGTCTCTC 1340
 Qy 1261 ACTGCTGACGGGAGGTGACAGAGAGAGCCCTGAGAACTCAAGCCGTGCTGATCTCA 1320
 Db 1341 ACTGCTGACGGGAGGTGACAGAGAGAGCCCTGAGAACTCAAGCCGTGCTGATCTCA 1400
 Qy 1321 GGGCAGCTGTCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1380
 Db 1401 GGGCAGCTGTCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1460
 Qy 1381 CAGCGTGAATCTGTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1440
 Db 1461 CAGCGTGAATCTGTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1520
 Qy 1441 GGCAGTGAATCTGTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1500
 Db 1521 GGCAGTGAATCTGTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 1580
 Qy 1501 CTCTCTCTCCACCGCAAGGGGCAATTCCTCAATGAGGAGTCTCCGCAAGGCTTA 1560
 Db 1581 CTCTCTCTCCACCGCAAGGGGCAATTCCTCAATGAGGAGTCTCCGCAAGGCTTA 1640
 Qy 1561 GAAGGCACTACCTCTGAGCACTTTGGCTCTCTGAGCACTGAGCTCTCTCTCTCTGAG 1620
 Db 1641 GAAGGCACTACCTCTGAGCACTTTGGCTCTCTGAGCACTGAGCTCTCTCTCTCTGAG 1700
 Qy 1621 GCGCGGCGCCAGCGCCCTCTGAGGGGCTGTGAGTGAAGACAGATCTCTCTCTCTGAGTCC 1680
 Db 1701 GCGCGGCGCCAGCGCCCTCTGAGGGGCTGTGAGTGAAGACAGATCTCTCTCTCTGAGTCC 1760
 Qy 1681 TTTGACCAATTTGACCTTGTGCTGAGCTTCCGAAACCCCACTGAGGGGCTGTGTGTCT 1740
 Db 1761 TTTGACCAATTTGACCTTGTGCTGAGCTTCCGAAACCCCACTGAGGGGCTGTGTGTCT 1820
 Qy 1741 GTGACCAACTGAGGGGCTTGAAGCACTCTCTCTGAGAGGTCTGAAGGATGTGTGGCAG 1800
 Db 1821 GTGACCAACTGAGGGGCTTGAAGCACTCTCTCTGAGAGGTCTGAAGGATGTGTGGCAG 1880

Qy 1801 GAATCTTGGGGGATAGCTGCTTTTCTTGAACAGATGCCAGAGGTGACTGCAAGCTTAC 1860
 Db 1881 GAATCTTGGGGGATAGCTGCTTTTCTTGAACAGATGCCAGAGGTGACTGCAAGCTTAC 1940
 Qy 1861 AGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
 Db 1941 AGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1973

RESULT 5
 BC046833
 LOCUS
 DEFINITION
 Mus musculus RIKEN cDNA 120013B22 gene, mRNA (CDNA clone MGC:61244
 IMAGE:6816569), complete cds.
 ACCESSION
 BC046833
 VERSION
 BC046833.1 GI:29145027
 KEYWORDS
 MGC.
 SOURCE
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Burow K.H., Scheaffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J., Mckernan K.J., Malek J.A., Gunaratne P.H., Richards S., Wierley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko I., Bouffard G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalske V., Smallus D.E., Scherch A., Schein J.F., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 2938)
 Strausberg R.
 Direct Submission
 Submitted (13-FEB-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contract: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas V. Casavant.

Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo M.F., Akabogu I., Bair T., Bair J., Crouch C., Davis A.,
 Fishler K., Kappel C., Kucaba T., Lebeck M., Melo A., Schaefer K.,
 Scheetz T., Smith C., Snir E., Tack D., Trout K., Walters J.,
 Casavant T., Soares M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 27229181.

FEATURES
source

gene

CDS

misc_feature

ORIGIN

Query Match 98.2%; Score 1859; DB 10; Length 2938;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

/gene="1200013822R1k"
/note="S. Tkc. Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain. Sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"
/db_xref="CDD:cd00180"

1 ATGAGTCCGTGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCGCTGAGC 60
81 ATGAGTCCGTGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCGCTGAGC 140
61 TCGAGAGCGCCGCTGCGGCGAGCGGCTCATCAAGTCGCTAAACCTCGAAGAAG 120
141 TCGAGAGCGCCGCTGCGGCGAGCGGCTCATCAAGTCGCTAAACCTCGAAGAAG 200
121 AAGCAGGCGGCTGAAGCGGCACTCAACAACAACCTGCGGCAAGCTACGAGTCCG 180
201 AAGCAGGCGGCTGAAGCGGCACTCAACAACAACCTGCGGCAAGCTACGAGTCCG 260
181 GAGACGCTGGCAAGGCACTCAAGGAAAGTGAAGAAGCAAGAGAGCTCGGGCGT 240
261 GAGACGCTGGCAAGGCACTCAAGGAAAGTGAAGAAGCAAGAGAGCTCGGGCGT 320
241 CTGAGTCCATCAAGTCCATCAAGAAAGCAAAATCAAGATAGCAAGATCTGTCGAC 300
321 CTGAGTCCATCAAGTCCATCAAGAAAGCAAAATCAAGATAGCAAGATCTGTCGAC 380

301 ATACGAGGAGATTTGAGATCATGTCTTACCTCAACCAACCCCAATCATTTGCATTCAT 360
381 ATACGAGGAGATTTGAGATCATGTCTTACCTCAACCAACCCCAATCATTTGCATTCAT 440
361 GAA-----GTTTGAAGATAGGAGCAAGATTGATTTGTC 396
441 GAAAGTGGGAGATCAACGCTTAAGTACGTGTGTAAGATAGGAGCAAGATTGATTTGTC 500
397 ATGAGTATCCAGCGCGAGCGCATCTGTATGATTAATCATCATGTAAGCGGCAAGCTGAGT 456
501 ATGAGTATCCAGCGCGAGCGCATCTGTATGATTAATCATCATGTAAGCGGCAAGCTGAGT 560
457 GAGCGGAGCGCAGGCAATTTCTTCCGACGATGTGTCTGCTCGTCACTACTGCCACG 516
561 GAGCGGAGCGCAGGCAATTTCTTCCGACGATGTGTCTGCTCGTCACTACTGCCACG 620
517 AACGGATCGTTCACCGAGATCTCAAGCTGGAAGAAACATCTTCTGATGCCAATGGAAC 576
621 AACGGATCGTTCACCGAGATCTCAAGCTGGAAGAAACATCTTCTGATGCCAATGGAAC 680
577 ATCAAGATTGCTGACTTTGGCTCTTCAACCTGTACCAAAAGCAAGTTCTCCAGAG 636
681 ATCAAGATTGCTGACTTTGGCTCTTCAACCTGTACCAAAAGCAAGTTCTCCAGAG 740
637 TTCTGTGGAGCCCTCTCTTACGCTTCCGCTGAGATAGTCAACGGAAAGCTTATGTGGC 696
741 TTCTGTGGAGCCCTCTCTTACGCTTCCGCTGAGATAGTCAACGGAAAGCTTATGTGGC 800
697 CCAGAGTGAAGAGCTGTCTGTGGCGCTTCTCTGTATATCTGTGTGATGGCAACAG 756
801 CCAGAGTGAAGAGCTGTCTGTGGCGCTTCTCTGTATATCTGTGTGATGGCAACAG 860
757 CCCTTGAAGGAGGATCATTAACACATGTGAAGCAATCATGTAACGGGCTTACCGT 816
861 CCCTTGAAGGAGGATCATTAACACATGTGAAGCAATCATGTAACGGGCTTACCGT 920
817 GAGCGGCCCAAGCCGTCGATGCTGTGGCTGTATCCGCTGTGTAATGTGAACCCC 876
921 GAGCGGCCCAAGCCGTCGATGCTGTGGCTGTATCCGCTGTGTAATGTGAACCCC 980
877 ACCCGTGGGCAACACTGAGAGATGTAGCCAGTCAATGGTGGTCAACTGGGTTACCC 936
981 ACCCGTGGGCAACACTGAGAGATGTAGCCAGTCAATGGTGGTCAACTGGGTTACCC 1040
937 ACCGAGTGGGGAACAGGAGCCCTGAGTGGGAGGAGGAGCACTGTGTGTGACTTTGGC 996
1041 ACCGAGTGGGGAACAGGAGCCCTGAGTGGGAGGAGGAGGAGCACTGTGTGTGACTTTGGC 1100
997 CGGCGCTCCATGCGGAGCTGTATAGTCTGCTCTCTGCGCCCTCTCTGAGAGATGAGCC 1056
1101 CGGCGCTCCATGCGGAGCTGTATAGTCTGCTCTCTGCGCCCTCTCTGAGAGATGAGCC 1160
1057 AAGGTGTGAGCTTCTTCAAGCAAGAGCTGCGGAGGTGAAGCACTGTACTGTGGCTG 1116
1161 AAGGTGTGAGCTTCTTCAAGCAAGAGCTGCGGAGGTGAAGCACTGTACTGTGGCTG 1220
1117 GAGCGGCAACATCTCTTGAAGAGTCCGGAAGGAGAAAGATAGCTCAAAATCTGCA 1176
1221 GAGCGGCAACATCTCTTGAAGAGTCCGGAAGGAGAAAGATAGCTCAAAATCTGCA 1280
1177 GGTGACCGGCTGAGAGATCTCTTCTGCGCTGCGCAAGAGAGCTTAAGCTTCCGAA 1236
1281 GGTGACCGGCTGAGAGATCTCTTCTGCGCTGCGCAAGAGAGCTTAAGCTTCCGAA 1340
1237 GGCATTCTCAAGAAAGTCTTCTTACCTGTCAAGGAGGTACAGAGAGCTTCAAGAA 1296
1341 GGCATTCTCAAGAAAGTCTTCTTACCTGTCAAGGAGGTACAGAGAGCTTCAAGAA 1400
1297 CTCAAGCGGAGTCTGATCTCAAGGAGCGCTGCTGCTGATACCTGTGCTCCAGAG 1356
1401 CTCAAGCGGAGTCTGATCTCAAGGAGCGCTGCTGCTGATACCTGTGCTCCAGAG 1460
1357 AAGGATCTCTTAAGAGTCTGAGACAGCTGAATCTGTACTTCTCTCAAGAGCC 1416

|||||
Db 1461 AAGGATCCTTAAAGATCTGACAGCGTGAATCTGTACTACTCTCTCCAGAGCCC 1520
|||
Qy 1417 AGGAGTCTGGGAACTCTTAAGACGCCAGTATGTGTTGTAGTGGGACCCGTGGAG 1476
|||
Db 1521 AGGAGTCTGGGAACTCTTAAGACGCCAGTATGTGTTGTAGTGGGACCCGTGGAG 1580
|||
Qy 1477 CAGAAGTCTCAGAGGCTTCAAGGCTCTCTCCACCGCAAGGGCANTCTCAATCAAT 1536
|||
Db 1581 CAGAAGTCTCAGAGGCTTCAAGGCTCTCTCCACCGCAAGGGCANTCTCAATCAAT 1640
|||
Qy 1537 GCGAAGTCTCCGCAAGCCTTGAAGGCACTACCCCTAGACCTTTGGCTCCCTGGAC 1596
|||
Db 1641 GCGAAGTCTCCGCAAGCCTTGAAGGCACTACCCCTAGACCTTTGGCTCCCTGGAC 1700
|||
Qy 1597 CAACTGGCTCTCTCCATCTCTGACGCCCGGCCAGCCCTCAGAGGGCTGTAGTAG 1656
|||
Db 1701 CAACTGGCTCTCTCCATCTCTGACGCCCGGCCAGCCCTCAGAGGGCTGTAGTAG 1760
|||
Qy 1657 GACAGATCTCTGCTCCGAGTCTTGAACCAATGGAATGCTGCTGAACGCTTCCGGA 1716
|||
Db 1761 GACAGATCTCTGCTCCGAGTCTTGAACCAATGGAATGCTGCTGAACGCTTCCGGA 1820
|||
Qy 1717 ACCCACTGAGGGGCTGTGTCTGTGACCAACTGAGGGGCTTGAAGACCTCCCTCA 1776
|||
Db 1821 ACCCACTGAGGGGCTGTGTCTGTGACCAACTGAGGGGCTTGAAGACCTCCCTCA 1880
|||
Qy 1777 GAAAGTCTGAAGGATGTGTGAGGGAATCTTGGGGGATGCTGCTTTCTTGAACAG 1836
|||
Db 1881 GAAAGTCTGAAGGATGTGTGAGGGAATCTTGGGGGATGCTGCTTTCTTGAACAG 1940
|||
Qy 1837 TGCCAAGAGGTGACTGACGCTTACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
|||
Db 1941 TGCCAAGAGGTGACTGACGCTTACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGC 1997
|||
RESULT 6
BC081899 2900 bp mRNA linear ROD 02-SEP-2004
LOCUS Rattus norvegicus cDNA clone MGC:93817 IMAGE:7110518, complete cds.
DEFINITION BC081899
ACCESSION BC081899.1 GI:51859485
VERSION MGC.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2900)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Mak, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Frange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,
Abrahamson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boulfard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
Schnerch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
human and mouse cdna sequences
Generation and initial analysis of more than 15,000 full-length
human and mouse cdna sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL PUMED 12477932
PUBMED 2 (bases 1 to 2900)
REFERENCE 12477932
AUTHORS Director MGC Project.

TITLE Direct Submission
JOURNAL Submitted (01-SEP-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-sngc.stanford.edu>
Contact: (Dickson, Mark) md@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAP Plate: 182 Row: a Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2900
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93817 IMAGE:7110518"
/issue_type="Kidney, rat (Brown Norway)"
/clone_1ib="NIH MGC_235"
/lab_host="DH10B"
/note="Vector: pEXpress1"
73..1965
/codon_start=1
/product="unknown (protein for MGC:93817)"
/protein_id="AAH81899.1"
/db_xref="GI:51859486"
/translation="MESVALHRGKGLVAPASALATESARPLADRLIKSPKPMKQAV
KHHHKNLRRHREPLETGGTGVKARSSGGLVAKISKDKIDBDLHHR
REILMSLNHPHITAIHEHPENSSKIVIMVAYRGDIYDYSRPRINERDAHFP
ROIYSAHCHONGIIVHDKENILILASNTIKIADPRLSYHKGRFLQTFCSPL
YASPEVNGKPYVGPVDSWSLGLVLYLVHGTMPDQDHTLVQISSGAIVRECK
PSDAGLIRMLMNVNIRATLDEVASHVMVMVGYSTRIGEBALREGHPSGSGRA
SMADWLRSSRPLLENKAVCSFPKQHYVPGSGTSGLEROSLKSRKEMQAQTLQ
NDPARDTSRPSKSLKLPKGIKKKASPSGVEGPOELPVSNTPOQVPAYPIIL
PRKGIKKSRROESGYVSPSPESGELLDADVDVSDPMEOKSPQASGRHARGIL
KNGFSTRALGSTRPSTFGSLDQLASHPPTRARARPSGAVESDILSSESDQDLP
ERLPETPLRGCVSVNLRLEQPPSEGLKRWQESLGDSFSLTDCQEVTAAYRQALG
ISKLS"
ORIGIN
Query Match 89.1%; Score 1686.6; DB 10; Length 2900;
Best local Similarity 93.6%; Pred. No. 0;
Matches 1771; Conservative 0; Mismatches 119; Indels 3; Gaps 1;
Qy 1 ATGAGTGGTGGCTTACTTCAAGCGCCGAGCAAGGCTCCCTGGGCTCCGCGCTGGCC 60
|||
Db 73 ATGAGTGGTGGCTTACTTCAAGCGCCGAGCAAGGCTCCCTGGGCTCCGCGCTGGCC 132
|||
Qy 61 TGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTGGCTTAACTGTATGAAG 120
|||
Db 133 ACGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTGGCTTAACTGTATGAAG 192
|||
Qy 121 AAGCAGCGCGGTAAACCGGACCATCAAAACAACCTCGCGCACCGCTACAGATTCCTG 180
|||
Db 193 AAGCAGCGCGGTAAACCGGACCATCAAAACAACCTCGGACACCGCTACAGATTCCTG 252
|||
Qy 181 GAGAGCTGGGCAAGGCGCACTACGAGGAAAGTGAAGAGGCAAGAGAGCTCGGGGCGT 240
|||
Db 253 GAGAGCTGGGCAAGGCGCACTACGAGGAAAGTGAAGAGGCAAGAGAGCTCGGGGCGT 312
|||

QY 241 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATATCAAGATGACAGATCTGCTGAC 300
 DB 313 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATATCAAGATGACAGATCTGCTGAC 372
 QY 301 ATACGAGGAGATGATGATCATGTCTTCACTCAACACCCCAATCATTTGCAATCAT 360
 DB 373 ATAGGAGGAGATGATGATCATGTCTTCACTCAACACCCCAATCATTTGCAATCAT 432
 QY 361 GAAAGTTTGAAGATGACAGCAAGATTTGATTTGATGATGATGATGATGATGATGAT 420
 DB 433 GAAAGTTTGAAGATGACAGCAAGATTTGATTTGATGATGATGATGATGATGATGAT 492
 QY 421 CTGATGATTTAATCAATGATGACGCGCAAGCTGATGATGATGATGATGATGATGAT 480
 DB 493 CTGATGATTTAATCAATGATGACGCGCAAGCTGATGATGATGATGATGATGATGAT 552
 QY 481 CGACAGATCGTGTCTGCGCCCTGCACTACTGCAACAGAACGAGATTTGTCACCGGACCTC 540
 DB 553 CGACAGATCGTGTCTGCGCCCTGCACTACTGCAACAGAACGAGATTTGTCACCGGACCTC 612
 QY 541 AAGCTGAAAACATCTCTTCTAGATGCAATGAAACATCAAGATTTGCTGATCTTTGCGCTC 600
 DB 613 AAGCTGAAAACATCTCTTCTAGATGCAATGAAACATCAAGATTTGCTGATCTTTGCGCTC 672
 QY 601 TCCACCTGTATCAACAAAGGCAAGTTCTTCAGACGTTCTGTGGAGCCCTCTCTAGCC 660
 DB 673 TCCACCTGTATCAACAAAGGCAAGTTCTTCAGACGTTCTGTGGAGCCCTCTCTAGCC 732
 QY 661 TCGCCTGAGATATCAACGCGGAAAGCCCTATGTGGGCCCAAGGTTGACAGCTGTCTGTG 720
 DB 733 TCACTGAGATATCAACGCGGAAAGCCCTATGTGGGCCCAAGGTTGACAGCTGTCTGTG 792
 QY 721 GAGGTTCTCTGTATCACTCTGTGATGCAATGCAACATGCGCTTTGACGCGGAGATCATATA 780
 DB 793 GAGGTTCTCTGTATCACTCTGTGATGCAATGCAACATGCGCTTTGACGCGGAGATCATATA 852
 QY 781 ACACTGATGAGCAAAATCAAGTAAACGCGGCTTACCGTGAAGCGCGCAAGCGTCCGATGCC 840
 DB 853 ACCCTGATGAGCAAAATCAAGTAAACGCGGCTTACCGTGAAGCGCGCAAGCGTCCGATGCC 912
 QY 841 TGTGGCCTGATCCGCTGTGCTGTTAATGTGTAACCCACCCGCTGGGCACTGAGAGAT 900
 DB 913 TGTGGCCTGATCCGCTGTGCTGTTAATGTGTAACCCACCCGCTGGGCACTGAGAGAT 972
 QY 901 GTAGCAGATCATGTGGGCTCAACTGCGGCTTACCAACCGGAGTCCGCGGAAACGAGAACCC 960
 DB 973 GTAGCAGATCATGTGGGCTCAACTGCGGCTTACCAACCGGAGTCCGCGGAAACGAGAACCC 1032
 QY 961 CTGCTGAGAGGATGAGGCACTCTAGTGTGATCTTTGCGCGGCTTCATGCGGATCTGCTTA 1020
 DB 1033 CTGCTGAGAGGATGAGGCACTCTAGTGTGATCTTTGCGCGGCTTCATGCGGATCTGCTTA 1092
 QY 1021 CGTGGCTCTCTGCGCCCTCTCTGAGATGAGCCAAAGTGTGACGCTTTCTTCAAGCAG 1080
 DB 1093 CGTGGCTCTCTGCGCCCTCTCTGAGATGAGCCAAAGTGTGACGCTTTCTTCAAGCAG 1152
 QY 1081 CAGGTGCGGAGATGAGGCACTGTACTGTGGCTGAGAGGCGCAACTTTCTTTAAGAG 1140
 DB 1153 CAGGTGCGGAGATGAGGCACTGTACTGTGGCTGAGAGGCGCAACTTTCTTTAAGAG 1212
 QY 1141 TCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGATGACCGGCTGAGATTAACCTCT 1200
 DB 1213 TCCGAAAGAGATGATGATGCTCAAAATCTGCAAGGATGACCGGCTGAGATTAACCTCT 1272
 QY 1201 TCTGCGCTCTGCAAGAGACCTTAAGCTTCCGAAAGGATTTCTCAAGAAAGTCTCTCT 1260
 DB 1273 TCTGCGCTCTGCAAGAGACCTTAAGCTTCCGAAAGGATTTCTCAAGAAAGTCTCTCT 1332
 QY 1261 ACCTGCTGAGGAGATGAGGCACTCTCAAGAACTCAAGACCGGCTGTGATCTCA 1320
 DB 1333 CCTCTCATGCGGAGATGAGGCACTCTCAAGAACTCAAGACCGGCTGTGATCTCA 1392
 QY 1321 GGGCAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAAAGCATCTTAAAGATCTCA 1380

DB 1393 GGGCAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAGGCAATCTTAAAGATCTCGG 1452
 QY 1381 CAGCTGAAATCTGCTTACTACTCTCTCCAGAGCCCAAGCTGTGGGAACTTTAGAC 1440
 DB 1453 CAGCTGAAATCTGCTTACTACTCTCTCCAGAGCCCAAGTGTGGGAACTTTAGAC 1512
 QY 1441 GCCAGTATGTGTTGTGATGAGGAGACCCGCTGAGAGAGATCTCCACAGCTTCAGG 1500
 DB 1513 GCAGTGTATGTGTTGTGATGAGGAGACCCCAAGAGAGATCTCCACAGCTTCAGG 1572
 QY 1501 CTCTCTCTCAAGGCAAGGCAATTTCAACTCAATGTGAGATTTCTCCGCAAGCTTTA 1560
 DB 1573 ---CGCTCATGTCAGAGGCAATCTCAACTCAATGTGAGATTTCTCCGCAAGCTTTA 1629
 QY 1561 GAAAGCATACCCCTGACACTTTGAGCTCTGAGACCAATGAGCTCTCCATCTGCA 1620
 DB 1630 GAAAGCATACCCCTGACACTTTGAGCTCTGAGACCAATGAGCTCTCCATCTGCA 1689
 QY 1621 GCCCGGCCAGCCGCTCTGAGGAGCTGTGATGAGAGCAGATCTGTCCTCGAGTCC 1680
 DB 1690 GCCCGGCCAGCCGCTCTGAGGAGCTGTGATGAGAGCAGATCTGTCCTCGAGTCC 1749
 QY 1681 TTTGACCAATTGGAATTTGCTGTAACGTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1740
 DB 1750 TTTGACCAATTGGAATTTGCTGTAACGTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1809
 QY 1741 GTGGAACAACCTGAGGGGCTTTGAGCAGCCCTCCCTGAGAGATGAGTGTGAGAGTGTG 1800
 DB 1810 GTGGAACAACCTGAGGGGCTTTGAGCAGCCCTCCCTGAGAGATGAGTGTGAGAGTGTG 1869
 QY 1801 GAATCTTGGGAGATGAGTCTTTCTCTGACAGACTGCAAGAGTGTGAGAGTGTGAG 1860
 DB 1870 GAATCTTGGGAGATGAGTCTTTCTCTGACAGACTGCAAGAGTGTGAGAGTGTGAG 1929
 QY 1861 AGACAAAGCCTAGGAATCTGCTCAAAAGCTCAGC 1893
 DB 1930 AGACAAAGCCTAGGAATCTGCTCAAAAGCTCAGC 1962
 RESULT 7
 AX380986
 LOCUS AX380986 2026 bp DNA linear PAT 18-MAR-2002
 DEFINITION Sequence 27 from Patent WO0212456.
 ACCESSION AX380986
 VERSION AX380986.1 GI:19575826
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus sp.
 Rattus sp.
 Buthyria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1
 Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.
 Ampk-related serine/threonine kinase, designated snark
 Patent: WO 0212456-A 27 14-FEB-2002;
 1149336 ONTRARIO INC. (CA)
 FEATURES
 source
 1..2026
 location/Qualifiers
 /organism="Rattus sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:10118"
 ORIGIN
 Query Match 88.9%; Score 1683.4; DB 6; Length 2026;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
 QY 1 ATGAGATCGTGGCTTACTCTCAAGCGCCGAGCAGAGCTCCCTCGGCTCCGCGCTGGCC 60
 DB 75 ATGAGATCGTGGCTTACTCTCAAGCGCCGAGCAGAGCTCCCTCGGCTCCGCGCTGGCC 134
 QY 61 TCGAGAGCGCCGCGCTGCGAGAGGCTCATCAAGTGGCTTAAAGCTGTGATGAG 120

Db 135 ACCGAGAGCGCCGCGCTGGCGGACCGGCTCATGAATGCGCCAAACCTCTGATGAAG 134
QY 121 AAGCAGCGCGTGAAGCGGACCATCAAAACAACCTGCGGCAACCGCTACGAGTTCTTG 180
Db 195 AAGCAGCGCGTGAAGCGGACCATCAAAACAACCTGAGGACCGCTACGAGTTCTTG 254
QY 181 GAGACGCTGGGCAAGGACCTACGGGAGGTGAAGAGGACGAGAGAGCTGGGGCGT 240
Db 255 GAGACCTGGGCAAGGACCTACGGGAGGTGAAGAGGAGAGAGCTGGGACGC 314
QY 241 CTGTGGCCCATCAAGTTCATCAGAAAAGCAAAATCAAGATGAGAGATCTGTCAC 300
Db 315 CTGTGGCCCATCAAGTTCATCAGAAAAGCAAAATCAAGATGAGAGATCTGTCAC 374
QY 301 ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCCCAATCATTTGCCATTCAT 360
Db 375 ATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCCCAATCATTTGCCATTCAT 434
QY 361 GAAGTCTTGAAGATGAGCAAGATTTGATTTGTCAATGAGTATGCGACGCGAGCGAT 420
Db 435 GAAGTCTTGAAGATGAGCAAGATTTGATTTGTCAATGAGTATGCGACGCGAGCGAT 494
QY 421 CTGTATGATTATCATAGTGAAGCGGCAACGCTGAGTGAAGCGGACGCGACGATTTCTTC 480
Db 495 CTGTATGATTATCATAGTGAAGCGGCAACGCTGAGTGAAGCGGACGCGACGATTTCTTC 554
QY 481 CGACATATGCTGTCTGCGCTGCACTACTGCGACCAAGCGGATCTTCAACGAGATCTC 540
Db 555 CGACATATGCTGTCTGCGCTGCACTACTGCGACCAAGCGGATCTTCAACGAGATCTC 614
QY 541 AAGCTGAAAACATCTCTTGAATGCGCAATGCAAGAAATCAAGATTTGCTGCTTTGCGCTC 600
Db 615 AAGCTGAAAACATCTCTTGAATGCGCAATGCAAGAAATCAAGATTTGCTGCTTTGCGCTC 674
QY 601 TCCAACTGTATCCAAAGGCAAGTCTTCCAGACGTTCTGTGAGAGCCCTCTCTACGCGC 660
Db 675 TCCAACTGTATCCAAAGGCAAGTCTTCCAGACGTTCTGTGAGAGCCCTCTCTACGCGC 734
QY 661 TCGCCCTGAATATCATACGGGAAAGCCCTATGTGGGCCCAAGGTTGAACAGTGTCTCTG 720
Db 735 TCACTGAAATCTTCAACGGGAAAGCCCTATGTGGGCCCAAGGTTGAACAGTGTCTCTG 794
QY 721 GGGCTCTCTGTATCATCTGATGCGCAATGCGACCATGCGCTTTGACGGGAGATCAATTA 780
Db 795 GGGCTCTCTGTATCATCTGATGCGCAATGCGACCATGCGCTTTGACGGGAGATCAATTA 854
QY 781 ACACTGTGAAGCAATCAATGCAAGGAGCTTACCGTGAAGCGGCAAGCGCTCGATGCGC 840
Db 855 ACCCTGTGAAGCAATCAATGCAAGGAGCTTACCGTGAAGCGGCAAGCGCTCGATGCGC 914
QY 841 TGTGAGCTGTATCGGTGGCTGTTAATGTGAACCCCAACCGTGGGCGACATGAGAGAT 900
Db 915 TGTGAGCTGTATCGGTGGCTGTTAATGTGAACCCCAACCGTGGGCGACATGAGAGAT 974
QY 901 GTAGCAGATCATTTGGGTGCACTGAGGAGTTTACCAACCGGATCGGGGAAACGAGAACGC 960
Db 975 GTAGCAGATCATTTGGGTGCACTGAGGAGTTTACCAACCGGATCGGGGAAACGAGAACGC 1034
QY 961 CTGCGTGAAGGAGTGAAGTGAATTTGAGCGGAGCTTCAATGAGCGAGTGAAGTGA 1020
Db 1035 CTGCGTGAAGGAGTGAAGTGAATTTGAGCGGAGCTTCAATGAGCGAGTGAAGTGA 1094
QY 1021 CGTGGCTCTCTCGCGCCCTCTCTGAGAAATGAGCCAAAGTGTGACGCTTTCAAGCAG 1080
Db 1095 CGTGGCTCTCTCGCGCCCTCTCTGAGAAATGAGCCAAAGTGTGACGCTTTCAAGCAG 1154
QY 1081 CAGGTGCGGAGGTGAAGCACTGTAACCTGGCTGAGAGGCAACATTTCTTAAGAG 1140
Db 1155 CAGGTGCGGAGGTGAAGCACTGTAACCTGGCTGAGAGGCAACATTTCTTAAGAG 1214
QY 1141 TCCCGAAAGAGATGATCATGTGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTCT 1200

Db 1215 TCCCGAAAGAGATGATCATGTGCTCAGACTCTGCAAGATGAGCCAGTTGAAGTACTTCC 1274
QY 1201 TCTGCGCCCTGGCAAGGACGAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAAGTCTCT 1260
Db 1275 TCTGCGCCCTGGCAAGGACGAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAAGTCTCT 1334
QY 1261 ACTGCTCAGAGGAGGTGAAGAGAGCCCTCAGAGAACTGAGACCGGTGCTGTAATCTCA 1320
Db 1335 CCTCATCGGGGAGGTGAAGAGAGCCCTCAGAGAACTGAGACCGGTGCTGTAATCTCA 1394
QY 1321 GGGCAGCTGTCTGCTGCTATCTCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTCGA 1380
Db 1395 GGGCAGCTGTCTGCTGCTATCTCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTCGG 1454
QY 1381 CAGCTGAATCTGTGTTACTACTCTCTTCCAGAGCCCAAGAGAGTCTGGGAACTCTTAAG 1440
Db 1455 CAGCTGAATCTGTGTTACTACTCTCTTCCAGAGCCCAAGAGAGTCTGGGAACTCTTAAG 1514
QY 1441 GCCAGTATGTGTTTGTGAGTGGGAGCCCGTGAAGCAGAAATCTCAAGAGCTTCAAGG 1500
Db 1515 GCGAGTATGTGTTTGTGAGTGGGAGCCCGTGAAGCAGAAATCTCAAGAGTCTCAAGG 1574
QY 1501 CTCCTCTCCACCGCAAGGCAATCTCAAACTCAATGAGCAAGTCTCCCGCAGAGCTTA 1560
Db 1575 ---CGCTCATCTGCAAGGCAATCTCAAACTCAATGAGCAAGTCTCCCGCAGAGCTTA 1631
QY 1561 GAAAGCACTACCTCTAGACCTTTGGCTCCCTGAGACCAATGAGCTCTCCATCTCTGA 1620
Db 1632 GAAAGCACTACCTCTAGACCTTTGGCTCCCTGAGACCAATGAGCTCTCCATCTCTGA 1691
QY 1621 GCCCGGCGCAGCGCCCTCAGGGGCTGAGAGGAGCAGAGATCTGCTCCGAGTCC 1680
Db 1692 GCCCGGCGCAGCGCTCTCTGAGAGCTGAGTGAAGCAGAAATCTGCTCCGAGTCC 1751
QY 1681 TTTGACCAATTTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 1752 TTTGACCAATTTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1811
QY 1741 GTGAGCAACCTGAGAGGAGCTTGAAGAGCTCTCTGAGAGGCTGAGAGGAGTGTGAG 1800
Db 1812 GTGAGCAACCTGAGAGGAGCTTGAAGAGCTCTCTGAGAGGCTGAGAGGAGTGTGAG 1871
QY 1801 GAATCTTGGGGAGATGAGCTTTCTGAGAGAGTGCAGAGGAGTGAAGTGAAGCTGAC 1860
Db 1872 GAATCTTGGGGAGATGAGCTTTCTGAGAGAGTGCAGAGGAGTGAAGAGTGAAGCTGAC 1931
QY 1861 AGACAGCCCTAGAGATCTGCTCAAGAGCTCAGC 1893
Db 1932 AGACAGCCCTAGAGATCTGCTCAAGAGCTCAGC 1964

RESULT 8
AX380960
LOCUS AX380960 2929 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 1 from Patent WO0212456.
ACCESSION AX380960
VERSION AX380960.1 GI:19575800
KEYWORDS
SOURCE
Rattus sp.
ORGANISM Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
AUTHORS Drucker, D.J., Rosen, C.F. and Iefevre, D.L.
TITLE Ampk-related serine/threonine kinase, designated snark
JOURNAL Patent: WO 0212456-A 1 14-FEB-2002;
1149336 ONTARIO INC. (CA)
FEATURES
source
1..2929
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

Query Match 88.9%; Score 1683.4; DB 6; Length 2929;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

1 ATGAGTCTGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 60
 83 ATGAGTCTGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCTGGCC 142
 61 TGGAGAGGCGCCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAACTCTGATGAAG 120
 143 AGGAGAGGCGCCGCGCTGGCGAGCGGCTCATCAAGTGGCTTAACTCTGATGAAG 202
 121 AAGCAGCGGCTGAAGGCGGCACTCAAAACAACCTGGGAGCGGCTCAAGTTCCTG 180
 203 AAGCAGCGGCTGAAGGCGGCACTCAAAACAACCTGGGAGCGGCTCAAGTTCCTG 262
 181 GAGACGCTGGGCGAAGGCGCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 263 GAGACGCTGGGCGAAGGCGCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 322
 241 CTGGTGGCATCAAGTCTCATAGAGAAACAATAAGATGAGAGATCTGCTGAC 300
 323 CTGGTGGCATCAAGTCTCATAGAGAAACAATAAGATGAGAGATCTGCTGAC 382
 301 ATACGAGGAGATTAAGATCAATGCTTCACTCAACCCCACTCAATGCTGACATCC 360
 383 ATACGAGGAGATTAAGATCAATGCTTCACTCAACCCCACTCAATGCTGACATCC 442
 361 GAAAGTTTGAAGATGAGAGCAAGATTTGATGATGAGATGAGATGAGATGAGAT 420
 443 GAAAGTTTGAAGATGAGAGCAAGATTTGATGATGAGATGAGATGAGATGAGAT 502
 421 CTGTATGATTAATCAAGTGAAGCGGCGCAAGGCTGAGAGCGGCGCAATTTCTTC 480
 503 CTGTATGATTAATCAAGTGAAGCGGCGCAAGGCTGAGAGCGGCGCAATTTCTTC 562
 481 CGACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 563 CGACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 541 AAGCTGGAAGAAATCATCTTCTAGATGCAATGAGAAATCATCAAGTGGCTGCTTC 600
 623 AAGCTGGAAGAAATCATCTTCTAGATGCAATGAGAAATCATCAAGTGGCTGCTTC 682
 601 TCCAACTGTACCAAAAGGCAAGTCTTCAGACGTTCTGAGAGCGCTCTCAAGCC 660
 683 TCCAACTGTATCAAAAGGCAAGTCTTCAGACGTTCTGAGAGCGCTCTCAAGCC 742
 661 TCCCTGAGATGATCAACGAGAAAGCCCTTATGAGGCGCAGAGGCTGCTGCTGCT 720
 743 TCCCTGAGATGATCAACGAGAAAGCCCTTATGAGGCGCAGAGGCTGCTGCTGCT 802
 721 GGGGTTCTCTGATCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 803 GGGGTTCTCTGATCAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 862
 781 ACACTGTGAAGAAATCAAGTGAAGGAGGCTTACCGTGAAGCGGCGCAAGCGCTG 840
 863 ACCCTGTGAAGAAATCAAGTGAAGGAGGCTTACCGTGAAGCGGCGCAAGCGCTG 922
 841 TGTGCTCTGATCCGCTGCTGCTTAAATGATGAACCCCACTGGGCGCACTGAGAGAT 900
 923 TGTGCTCTGATCCGCTGCTGCTTAAATGATGAACCCCACTGGGCGCACTGAGAGAT 982
 901 GTAGCAGATCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
 983 GTAGCAGATCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1042
 961 CTGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
 1043 CTGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1102

1021 CGTCTGCTCTGCGGCGCCCTCTCTGAGAAATGAGCCAGGATGAGCTTCTTCAAGCAG 1080
 1103 CGTCTGCTCTGCGGCGCCCTCTCTCTGAGAAATGAGCCAGGAGGAGGAGGAGGAG 1162
 1081 CAGCTGCGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
 1163 CAGCTGCGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1222
 1141 TCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 1223 TCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1282
 1201 TCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
 1283 TCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1342
 1261 ACCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
 1343 CCTCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1402
 1321 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
 1403 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1462
 1381 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
 1463 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1522
 1441 GCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
 1523 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1582
 1501 CTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
 1583 ---CGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1639
 1561 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
 1640 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1699
 1621 GCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
 1700 GCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1759
 1681 TTTGACCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
 1760 TTTGACCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1819
 1741 GTGGAACAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
 1820 GTGGAACAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1879
 1801 GAATCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
 1880 GAATCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1939
 1861 AGACAAAGCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1893
 1940 AGACAAAGCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1972

RESULT 9
 AC125887 291762 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
 DEFINITION ***; 13 unordered pieces.
 AC125887
 ACCESSION AC125887.3 GI:25008671
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 291762)
Munz, D., Maric, M., Lee, A., Lee, A., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebech, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, T., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Huik, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, U.,
Lorenshew, L., Louie, H., Lozano, R.J., Lu, X., Ma, U.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawlin, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Mijic, B., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwagwu, O., Okunolu, G., Olamposagun, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Rutz, S.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smales, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Vais, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
2 (bases 1 to 291762)
Unpublished
Worley, K.C.
Direct Submission
Submitted (02-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 291762)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23269681.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMD
Center clone name: CH230-27473
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 230953 bases at least Q30
Consensus quality: 233437 bases at least Q20
Estimated insert size: 233445; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 70277: contig of 70277 bp in length
70278 70377: gap of unknown length
70378 95131: contig of 24654 bp in length
95032 108956: gap of unknown length
108957 109056: contig of 13825 bp in length
109057 271697: gap of unknown length
271698 271797: contig of 162641 bp in length
271798 273026: gap of unknown length
273027 273126: contig of 1229 bp in length
273127 274531: gap of unknown length
274532 274631: contig of 1405 bp in length
274632 276287: gap of unknown length
276288 276387: contig of 1656 bp in length
276388 277904: gap of unknown length
277905 278004: contig of 1517 bp in length
278005 279523: gap of unknown length
279524 279623: contig of 1519 bp in length
279624 280724: gap of unknown length
280725 280824: contig of 1101 bp in length
280825 284658: gap of unknown length
284659 284758: contig of 384 bp in length
284759 287234: gap of unknown length
287235 287334: contig of 246 bp in length
287335 291762: gap of unknown length
291762 4428: contig of 4428 bp in length.

Location/Qualifiers
1..291762
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-27473"
70378..75596
/note="wgs contig"
95132..96554
/note="wgs contig"
109057..110574
/note="wgs contig"

ORIGIN
Query Match 80.7%; Score 1528.4; DB 2; Length 291762;
Best Local Similarity 99.3%; Pred. No. 0;

Matches 1535; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGAGTCGGTGGCTTACTTACAGCGCCGAGCGCTCCTCGGCTCCGCTGGCC 60
DB 274742 ATGAGTCGGTGGCTTACTTACAGCGCCGAGCGCTCCTCGGCTCCGCTGGCC 274801

QY 61 TCGGAGAGCGCCCGCGCTGGCGAGCGGCTCATCAAGTCGCTTAACTCTGATGAG 120
DB 274802 TCGGAGAGCGCCCGCGCTGGCGAGCGGCTCATCAAGTCGCTTAACTCTGATGAG 274861

QY 121 AAGCAGCGCGTGAAGCGCGCATCAACAACAACCTGGGAGACCGCTCGAGTTCG 180
DB 274862 AAGCAGCGCGTGAAGCGCGCATCAACAACAACCTGGGAGACCGCTCGAGTTCG 274921

QY 181 GAGACGCTGGGCAAGGCGCATCAAGGAGAGTGAAGAAGGCAAGAGAGCTCGGCGCT 240
DB 274922 GAGACGCTGGGCAAGGCGCATCAAGGAGAGTGAAGAAGGCAAGAGAGCTCGGCGCT 274981

QY 241 CTGGTGGCCATCAAGTCATCAAGAAAGACAAATCAAGATGACAGATCTGCTGAC 300
DB 274982 CTGGTGGCCATCAAGTCATCAAGAAAGACAAATCAAGATGACAGATCTGCTGAC 275041

QY 301 ATACGAGAGGAGATGAGATCATGTCTTCACTGACCAACCCGACATCATTCGATCAT 360
DB 275042 ATACGAGAGGAGATGAGATCATGTCTTCACTGACCAACCCGACATCATTCGATCAT 275101

QY 361 GAAGTGTGAGATAGACAGCAAGATTTGATTTGTCAAGATGATGCGACCGAGCGAT 420
DB 275102 GAAGTGTGAGATAGACAGCAAGATTTGATTTGTCAAGATGATGCGACCGAGCGAT 275161

QY 421 CTGTATGATTAATCATAGTGAAGCGGCGCAAGGCTGAGAGCGGAGCGCAGGCAATTTCT 480
DB 275162 CTGTATGATTAATCATAGTGAAGCGGCGCAAGGCTGAGAGCGGAGCGCAGGCAATTTCT 275221

QY 481 CGACAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 275222 CGACAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275281

QY 541 AAGCTGAAAAAATCTCTTCTGATGAGCAATGGAATGGAATGCTGATGCTTGGGCTC 600
DB 275282 AAGCTGAAAAAATCTCTTCTGATGAGCAATGGAATGGAATGCTGATGCTTGGGCTC 275341

QY 601 TCCAACTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGAGGAGCCCTCTTACGCC 660
DB 275342 TCCAACTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGAGGAGCCCTCTTACGCC 275401

QY 661 TCGCTGAGATAGTCAACGAGGAGCCCTATGATGAGGCGGAGGAGTGAACAGCTGCTCTG 720
DB 275402 TCGCTGAGATAGTCAACGAGGAGCCCTATGATGAGGCGGAGGAGTGAACAGCTGCTCTG 275461

QY 721 GGGGTTCTCTGTACATCTCTGTGATGAGCAATGCGATGCGCTTTGACGGGAGGATCATTA 780
DB 275462 GGGGTTCTCTGTACATCTCTGTGATGAGCAATGCGATGCGCTTTGACGGGAGGATCATTA 275521

QY 781 ACACTGTGAGAGAAATCAATAGAGGAGCTTACCGTGAAGCGGCAAGCGCTCGAGTGC 840
DB 275522 ACACTGTGAGAGAAATCAATAGAGGAGCTTACCGTGAAGCGGCAAGCGCTCGAGTGC 275581

QY 841 TGGGCTGTATCCGGTGGCTGTTAATGATGAGACCCCAACCGCTGGGCGCACATGAGAT 900
DB 275582 TGGGCTGTATCCGGTGGCTGTTAATGATGAGACCCCAACCGCTGGGCGCACATGAGAT 275641

QY 901 GTAGCAGATATGATGAGGAGCACTGGGTTAACAACCGAGATCGGGGAAACGAGAACCC 960
DB 275642 GTAGCAGATATGATGAGGAGCACTGGGTTAACAACCGAGATCGGGGAAACGAGAACCC 275701

QY 961 CTGCGTGAAGGAGGAGCACTTGAATGATGAGCTTTGAGCGGCGCTCATGAGCGAGCTGTTA 1020
DB 275702 CTGCGTGAAGGAGGAGCACTTGAATGATGAGCTTTGAGCGGCGCTCATGAGCGAGCTGTTA 275761

QY 1021 CGTGGCTCTCTCGGCGCCCTCTCTGAGAGATGAGAGCAAGGTGTGACGCTTCTTCAAGCAG 1080
DB 275762 CGTGGCTCTCTCGGCGCCCTCTCTGAGAGATGAGAGCAAGGTGTGACGCTTCTTCAAGCAG 275821

QY 1081 CACGTGCGGAGAGTGAAGCACTGTACTTGGGCTGAGCGGCAATTCTTTAGAG 1140
DB 275822 CACGTGCGGAGAGTGAAGCACTGTACTTGGGCTGAGCGGCAATTCTTTAGAG 275881

QY 1141 TCCCGAAAGAGATGATAGTGGCTCAAAATCTGCAAGTGAACCGGCTGAGTACTCT 1200
DB 275882 TCCCGAAAGAGATGATAGTGGCTCAAAATCTGCAAGTGAACCGGCTGAGTACTCT 275941

QY 1201 TCTGCGCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTCT 1260
DB 275942 TCTGCGCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTCT 276001

QY 1261 ACCTGTGAGGAGGATGAGAGAGACCTTCAAGAACTCAAGACCGGCTGATATCTCA 1320
DB 276002 ACCTGTGAGGAGGATGAGAGAGACCTTCAAGAACTCAAGACCGGCTGATATCTCA 276061

QY 1321 GGGCAGCTGTCTCTGCTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 276062 GGGCAGCTGTCTCTGCTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276121

QY 1381 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCGAGCGAGTCTGGGAACTTTAGAC 1440
DB 276122 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCGAGCGAGTCTGGGAACTTTAGAC 276181

QY 1441 GCCAGTGAATGTTGTGATGAGGAGACCCGCTGAGAGCAGAACTCTCACAGGCTTCAGG 1500
DB 276182 GCCAGTGAATGTTGTGATGAGGAGACCCGCTGAGAGCAGAACTCTCACAGGCTTCAGG 276241

QY 1501 CTCCTCTCCACCGCAAGGCGATTCTGAACTCAATGCGAAGTTCT 1546
DB 276242 CTCCTCTCCACCGCAAGGCGATTCTGAACTCAATGCGAAGTTCT 276287

RESULT 10
AX407118 1884 bp DNA linear PAT 14-JUN-2002
LOCUS AX407118
DEFINITION Sequence 3 from Patent WO0224921.
ACCESSION AX407118
VERSION AX407118.1 GI:21439883
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Curtis, R.A. and Galvin, K.M.
AUTHORS 3700, a novel human protein kinase and uses therefor
TITLE Patent: WO 0224921-A 3 28-MAR-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..1884
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 71.9%; Score 1360.2; DB 6; Length 1884;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

QY 47 CCTCGGCGCTGGCTGAGAGCGCGCGCGCTGAGCGGCTCATCAAGTCGCTTA 106
DB 35 CCACTCTCTCGGCGCGAGAGCTAGCCCGCGCGCTGAGCGGCTGATTAAGTCGCCCA 94

QY 107 AACCTGTATGAAGAGAGCGCGGTGAAGCGGCAACATCAAAACAACCTGCGGACCC 166
DB 95 AGCCCTTAATGAAGAGAGCGCGGTGAAGCGGCAACATCAAAACAACCTGCGGACCC 154

QY 167 GCTACAGATTCCTGAGAGCGCTGGGCAAGGCACTTACGGAAGTGAAGAGGAGGAG 226
DB 155 GCTACAGATTCCTGAGAGCGCTGGGCAAGGCACTTACGGAAGTGAAGAGGAGGAG 214

227 AGAGCTCGGGGCGTCTGTGGCCATCAAGTCCATCGAAGAAAGCAAAATGCAAGATGAGC 286
215 AAGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCGAAGAAAGCAAAATGCAAGATGAGC 274
287 AGAGTCTGTGCAATACGAGGAGGAGATTGAGATCAATGTCTTCACTCAACACCCGAC 346
275 AAGATCTGATGCAATACGAGGAGGAGATTGAGATCAATGTCTTCACTCAACACCCGAC 334
347 TCATTGCCATCCATGAAAGTGTGAGAAATGCAAGCAAGATTGATGTGATGAGATG 406
335 TCATTGCCATCCATGAAAGTGTGAGAAATGCAAGCAAGATTGATGTGATGAGATG 394
407 CCAGCGAGGCGCATCTGTATGATTATCAATGAGCGGCGCATGAGTGTGAGCGGAG 466
395 CCAGCGGCGGCGCATCTGTATGATTATCAATGAGCGGCGCATGAGTGTGAGCGGAG 454
467 CCAGGCAATTTCTTCCGACAGATGTGTGCTGCTGCACTACTGCAAGAGGAGATG 526
455 CTAGGCAATTTCTTCCGACAGATGTGTGCTGCTGCACTACTGCAAGAGATG 514
527 TTCAACGAGATCTCAAGCTGAGAAACATCTTCAATGCGAATGGAACATCAAGATTG 586
515 TCACACGAGATCTCAAGCTGAGAAACATCTTCAATGCGAATGGAACATCAAGATTG 574
587 CTGACTTTGAGCTTCCAACTGTACCACTGACAAAGGCAAGTTCCCAAGAGTCTGTGG 646
575 CTGACTTTGAGCTTCCAACTGTACCACTGACAAAGGCAAGTTCCCAAGAGTCTGTGG 634
647 GCCCTCTCAACGCTGCGCTGAGATGATCAACGAGGAGCCCTTATGTGGCCAGAGTGG 706
635 GCCCTCTATGCTGCGGAGATTGTCAATGGAAGCCCTTACAGAGCCCAAGAGTGG 694
707 ACAAGCTGTCTCTGGGCGTTCTCTGTACATCTTGTGATGAGCAACATGCTTTGAGC 766
695 ACAAGCTGTCTCTGGGCGTTCTCTGTACATCTTGTGATGAGCAACATGCTTTGAGC 754
767 GGCAGGATCATATAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCCGCCA 826
755 GGCAGGATCATATAACACTGATGAGCAATCAATGAGGAGGCTTACCGGAGCCGCCA 814
827 AGCCGCTCGATGCTGTGCTGATCCGCTGCTGCTTAAATGTGAAACCCGACCCGCTGG 886
815 AACCCCTGATGCTGTGCTGATCCGCTGCTGCTTAAATGTGAAACCCGACCCGCTGG 874
887 CCAACTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 946
875 CCAACTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATG 934
947 GGGAAACAGGAGCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006
935 GAGAGCAGAGGCTCCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 994
1007 TGGCGGACTGTTCAGTCTGCTCTCTCGGCGCCCTCTCTGAGAGATGAGCCAGGTGCA 1066
995 TGGCGGACTGTTCAGTCTGCTCTCTCGGCGCCCTCTCTGAGAGATGAGCCAGGTGCA 1054
1067 GCTTCTTCAAGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1126
1055 GCTTCTTCAAGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1114
1127 ATTCTCTTAAAGTCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186
1115 ATTCTCTTAAAGTCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1174
1187 CTGAGGATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
1175 CTGAGGATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
1247 AAAAAAAGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1306
1235 AAAAAAAGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1294
1307 TGCCTGATATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1366

1295 TCCCTCGAGCCGAGGCGAGGCTGCCCC-----CCCTGCTCCCAAGAGGCAATTC 1345
1367 TTAAAGATCTCGACAGCGTGAATCTGTATCACTCTCTCGAAGCCGAGGATGCT 1426
1346 TCAGAGAGCCCGGACAGCCGAGGTGTGGCTACTACTCTCCGAGCCGAGGATGCTG 1405
1427 GGGAACTCTTAAAGCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1486
1406 GGGAGCTCTTGAACGAGGCGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1465
1487 CACAGGCTTCAAGGCTCTCTCTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1546
1466 CCGAGCTTCAAGGCTCTCTCTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1525
1547 CCGGACAGCCCTTGAAGGCACTACCCCTGAGCACTTTGCTCTCTGAGCACTGAGCT 1606
1526 CCGAGACAGCCCTTGAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1585
1607 CTTCCATCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
1586 CACTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1645
1667 TGTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1726
1646 TGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1705
1727 GGGGCTGT 1780
1706 GGGGCTGT 1765
1781 -----GTCTGAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1834
1766 GAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1825
1835 ACTGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1893
1826 ACTGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1884

RESULT 11
AX407116
LOCUS AX407116 3353 bp DNA linear PART 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0224921.
ACCESSION AX407116
VERSION AX407116.1 GI:21439882
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1
Curtis, R.A. and Galvin, K.M.
3700, a novel human protein kinase and uses therefor
Patent: WO 0224921-A 1 28-MAR-2002;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..3353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 71.9%; Score 1360.2; DB 6; Length 3353;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

47 CCTCGCCCTGCGCTTCGAGAGCGCCGCGCTGAGCGGAGCGGCTCATCAAGTGCCTTA 106
191 CCACTCCCTCGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 250
107 AACCTGTATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 166

Db 251 AGCCCTTAATGAAGAGAGGCGGTGAGCGGACCAACCAAGCAAACTGCGGACCC 310
 QY 167 GCTACAGATTCTCGAGAGCGTGGGCAAGGGCACTACGGGAAGTGAAGAGGACGAG 226
 Db 311 GCTACAGATTCTCGAGAGCGTGGGCAAGGGCACTACGGGAAGTGAAGAGGCGG 370
 QY 227 AGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAAATGAC 286
 Db 371 AGAGCTCGGGGCGCTGTGGCCATCAAGTCCATCAGAAAAGCAAAATCAAAATGAC 430
 QY 287 AGAGTGTCTGCAATACGAGGAGGAGTGAAGTCAATGCTTCACTCAACACCCCA 346
 Db 431 AAGATCTGATGCAATACGAGGAGGAGTGAAGTCAATGCTTCACTCAACACCCCA 490
 QY 347 TCATTCGCAATCCATGAAGTGTGAGAAATAGCAGAAAGTGTGATGTCTAGAGTATG 406
 Db 491 TCATTCGCAATCCATGAAGTGTGAGAAAGCAGCAAGATGCTGATGCTATGAGATATG 550
 QY 407 CCAAGCGAGGCGCATCTGATGATTAATCATAGAGCGGCGCAAGCTGATGAGCGGAGCG 466
 Db 551 CCAAGCGGCGGCGCATCTTATGATCATATCAGAGGCGGCGAGCTCAGTGAAGCGGAG 610
 QY 467 CCAAGCATTTCTTCCGACAGATGCTGTGCTGCTGCACTATGCTCAGCAAGAGGATG 526
 Db 611 CTAGGCAATTTCTTCCGAGAGATGCTGCTGCTGCACTATGCTCAGCAAGAGATG 670
 QY 527 TTCAACGAGATCTCAAGCTGGAAGAAAGATCTTATAGTGCATGGAATCAAGATTTG 586
 Db 671 TCACACGAGATCTCAAGCTGGAAGAAAGATCTTATAGTGCATGGAATCAAGATTTG 730
 QY 587 CTGACTTTGGCCCTCTCAACCTGTACCAACAAAGGCAAGTCTCTCAGAGCTTCTGAG 646
 Db 731 CTGACTTTGGCCCTCTCAACCTGTACCAACAAAGGCAAGTCTCTCAGAGCTTCTGAG 790
 QY 647 GCCCTCTTACAGCTCTGCTGAGATGATCAAGGAAAGCTTATGAGGCGGCGAGAGT 706
 Db 791 GCCCTCTTACAGCTCTGCTGAGATGATCAAGGAAAGCTTATGAGGCGGCGAGAGT 850
 QY 707 ACAGCTGTCTCTGAGGCGTCTCTGATCACTGTGATGATGAGCAACATGCGCTTGAAG 766
 Db 851 ACAGCTGTCTCTGAGGCGTCTCTGATCACTGTGATGATGAGCAACATGCGCTTGAAG 910
 QY 767 GCGAGATATTAACAAGCTGTGAGCAAAATCAGTAACGGGCTTACCGTGAAGCGGCA 826
 Db 911 GCGATGACCTAAGATCTTATGATGAACATGACAAAGGCGGCTTACCGGAGACCACTA 970
 QY 827 AGCCGCTCGATGCTGTGCTGATCGGTGCTGTTATGATGAAACCCCAACCCGCTG 886
 Db 971 AACCCCTCTGATGCTGTGCTGATCGGTGCTGTTATGATGAAACCCCAACCCGCTG 1030
 QY 887 CCAACAATTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATG 946
 Db 1031 CCAACCTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATG 1090
 QY 947 GGGAAAGAGAGGCTGCGTGAAGGAGGAGCAACCTGATGATGATGATGATGATGATG 1006
 Db 1091 GAGAGAGAGAGGCTGCGTGAAGGAGGAGCAACCTGATGATGATGATGATGATGATG 1150
 QY 1007 TGGCGAGCTGTGATGCTGCTCTGCGCCCTCTCTGAGAAATGAGCCCAAGGTGAC 1066
 Db 1151 TGGCTGATGCTGCTGCGGCGGCTCTCTCTCTCTGAGAAATGAGCCCAAGGTGAC 1210
 QY 1067 GCTTCTTCAAGAGCAAGTGGCGGAGGAGTGAAGCACTGATGATGATGATGATGATG 1126
 Db 1211 GCTTCTTCAAGAGCAAGTGGCGGAGGAGCAACCTGATGATGATGATGATGATG 1270
 QY 1127 ATTCTTGAAGATGCTGCGGAAAGAGATGATGATGATGATGATGATGATGATGATG 1186
 Db 1271 ATTCTTGAAGATGCTGCGGAAAGAGATGATGATGATGATGATGATGATGATGATG 1330
 QY 1187 CTGAGATGATGCTTCTGCGCTGCGAGAGCAAGCTTAAAGTTCGAAAGGCAATTTCT 1246
 Db 1331 CTGATGATGATGCTGCGCTGCGAGAGCAAGCTTAAAGTTCGAAAGGCAATTTCT 1390

QY 1247 AGAAAAAGTCTCTTACTTGTGAGGGAGATACAGAGGAGCCCTCAGAACTCAGACCG 1306
 Db 1391 AGAAGAGGATGTCAGGCTCTGCAAGAGGGGTACAGAGGAGCCCTCGAGCTCAGCCAA 1450
 QY 1307 TGGCTGATGATCTCAGAGGAGGAGGCTGTGCTGCTGATGATGATGATGATGATG 1366
 Db 1451 TGGCTGATGATCTCAGAGGAGGAGGCTGTGCTGCTGATGATGATGATGATGATG 1501
 QY 1367 TTAAGAGTGTGAGCAGCGTGAATGATGATGATGATGATGATGATGATGATGATG 1426
 Db 1502 TCAAGAGGCGGAGCAGCGGAGGAGTGTGATGATGATGATGATGATGATGATGATG 1561
 QY 1427 GGGAACTTTAAGACGCAAGTGTGATGATGATGATGATGATGATGATGATGATG 1486
 Db 1562 GGGAGCTTTGAGCGCAGGAGGAGGATGATGATGATGATGATGATGATGATGATG 1621
 QY 1487 CACAGGCTTCAAGGAGGCTCTCTCTCAGCAGAGGAGATTTCAAACTCAATGGAAT 1546
 Db 1622 CGCAAGCTTCAAGGAGGCTCTCTCTCAGCAGAGGAGATTTCAAACTCAATGGAAT 1681
 QY 1547 CCGCAGACGCTTGAAGAGGAGCAGTACCTTGAAGCCTTGAAGCCTTGAAGCAG 1606
 Db 1682 CCGCAGACGCTTGAAGAGGAGCAGTACCTTGAAGCCTTGAAGCCTTGAAGCAG 1741
 QY 1607 CCTCCATGCTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
 Db 1742 CACCTGCGCCCTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801
 QY 1667 TGTCTCCAGATCTTTGAACAATGGAATGATGATGATGATGATGATGATGATGAT 1726
 Db 1802 TGTCTCCAGATCTTTGAACAATGGAATGATGATGATGATGATGATGATGATGAT 1861
 QY 1727 GGGGCTGT 1780
 Db 1862 GGGGCTGT 1921
 QY 1781 -----GTCTGAAGCGATGATGATGATGATGATGATGATGATGATGATGATG 1834
 Db 1922 GAAAGCTGCTGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1981
 QY 1835 ACTGCAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1993
 Db 1982 ACTGCAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040

RESULT 12
 AX642961 3360 bp DNA linear PAT 24-FEB-2003
 LOCUS AX642961
 DEFINITION Sequence 38 from Patent WO01096547.
 ACCESSION AX642961
 VERSION AX642961.1 GI:28550104
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Yue, H., Lal, P., Bandman, O., Borowsky, M., Au-Young, J., Lu, Y.,
 Gandhi, A.R., Tribouley, C.M., Wallia, N., Yao, M.G., Lu, D.A.,
 Greenwald, S.R., Ramkumar, J., Griffin, J.A., Kearney, L., Buford, N.,
 Nguyen, D.B., Tang, Y.T., Baughn, M.R., He, A., Thornton, M.,
 Hafalia, A., Patterson, C., Gururajan, R., Lo, T.P., Khan, F.,
 Rejzbon, S.A., Azimzal, Y., Policky, J.L., Ding, L., Grether, M.,
 Elliott, V.S., Thangavelu, K., Batra, S. and Ison, C.H.
 Humain kinases
 Patent: WO 01096547-A 38 20-DEC-2001;
 Incyte Genomics, Inc. (US)
 TITLE
 JOURNAL
 FEATURES
 source
 1.3360
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN /note="Incyte ID No: 4841542GB1"

Query Match 71.7%; Score 1357.6; DB 6; Length 3360;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;

QY 32 GCCAGGCTCCCTGGGCTCCGCTCCGAGAGAGCCCGGCTGGCGGACGAGC 91
DB 98 GCGGCGGCTCCGAGCCCTCCCTCCGCGCAGAGCTAGCCCGGCTGGCGGAGGCG 157
QY 92 TCATCAAGTCGCTTAACCTCTGATGAAAGAGCGGCGGTGAAGCGGACCATCAAA 151
DB 158 TGAATCAAGTCGCCAAGCCCTTAATGAAGAGCGGCGGTGAAGCGGACCATCAAGC 217
QY 152 ACAACCTGCGGCAACCGCTACGAGTTCCTGAGAGCGCTGGCGCAAGGCACTTACGCGAAGG 211
DB 218 ACAACCTGCGGCAACCGCTACGAGTTCCTGAGAGCGCTGGCGCAAGGCACTTACGCGAAGG 277
QY 212 TGAAGAAAGCAAGAGAGCTCGGCGGCTGTGGTGGCATCAAGTCCATAGAGAAAGCA 271
DB 278 TGAAGAAAGCGCGGAGAGCTCGGCGGCTGTGGTGGCATCAAGTCCATAGAGAAAGCA 337
QY 272 AATCAAGATGAGAGATCTGCTGCAATACGAGAGAGATTGAGATCATGCTTAC 331
DB 338 AATCAAGATGAGAGATCTGCTGCAATACGAGAGAGATTGAGATCATGCTTAC 397
QY 332 TCAACCAACCCCAATCATTCATTCATGAGTGTGAAGATGAGAGAGATTGTA 391
DB 398 TCAACCAACCCCAATCATTCATTCATGAGTGTGAAGATGAGAGAGATTGTA 457
QY 392 TTGTCAAGATGAGAGAGAGAGAGAGATGATGATTAATCAATCAAGTGAAGCGCACGCG 451
DB 458 TGTCAAGATGAGAGAGAGAGAGAGATGATGATTAATCAATCAAGTGAAGCGCACGCG 517
QY 452 TGAAGAGAGAGAGAGAGAGATTTCTTCCGACAGATCGTGTGCGGCTGCAATGCGC 511
DB 518 TGAAGAGAGAGAGAGAGATTTCTTCCGACAGATCGTGTGCGGCTGCAATGCGC 577
QY 512 ACCAGAAAGGAGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTCTAGATGCCAATG 571
DB 578 ATCAGAAAGGAGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTCTAGATGCCAATG 637
QY 572 GAAACATCAAGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTCTAGATGCCAATG 631
DB 638 GGAATATCAAGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTCTAGATGCCAATG 697
QY 632 AGACGTTCTGTGGAGAGCCCTCTAGAGCCCTGAGATGAGATGAGAGAGAGCCCTATG 691
DB 698 AGACGTTCTGTGGAGAGCCCTCTAGAGCCCTGAGATGAGATGAGAGAGAGCCCTATG 757
QY 692 TGGGCGCAGAGTGAAGCTGTCTCTGGGCGTTCTCTGTATCATCTGTGATGATGCA 751
DB 758 CAGGCGCAGAGTGAAGCTGTCTCTGGGCGTTCTCTGTATCATCTGTGATGATGCA 817
QY 752 CCAATGCTTTTGAAGCGGCGAGATCAATAAACATCTGTGAAGCAATCAAGTGAAGCGGCGCTT 811
DB 818 CCAATGCTTTTGAAGCGGCGAGATCAATAAACATCTGTGAAGCAATCAAGTGAAGCGGCGCTT 877
QY 812 ACCGTGAGCGCGCAGAGCGCTCGATGCGCTGTGGCGGTGATCGGATGAGTGAATGATGA 871
DB 878 ACCGTGAGCGCGCAGAGCGCTCGATGCGCTGTGGCGGTGATCGGATGAGTGAATGATGA 937
QY 872 ACCCGACCGCTCGGCGCACAATGAGATGATGAGATGATGAGATGATGAGATGATGAGATG 931
DB 938 ACCCGACCGCTCGGCGCACAATGAGATGATGAGATGATGAGATGATGAGATGATGAGATG 997
QY 932 ACAACCAACCGAGTCTGGGAGAAAGAGAGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 991
DB 998 ACAACCAACCGAGTCTGGGAGAAAGAGAGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1057
QY 992 TTGGCGGCGCTCTCATGCGGAGATGTTGATGCTGCTCGGCGCGCTCTGAGAGAAAG 1051

DB 1058 CTGCCCCGCGCTTCATAGGCTGACTGCTCCGCGCTTCTCCGCCCTCTCTGAGAAATG 1117
QY 1052 GAGCCAAAGGTGTGACGCTTCTTCAAGACAGACGCTGCGGAGGTGAGAGCATGTACTG 1111
DB 1118 GGGCCAAAGGTGTGACGCTTCTTCAAGACAGACGCTGCGGAGGTGAGAGCATGTACTG 1177
QY 1112 GGTGTGAGGGGCAACATCTCTTAAGAAATGCCGAAGAGAAATGACATGCTCAAAATC 1171
DB 1178 GCTGTGAGGGGCAACATCTCTTAAGAAATGCCGAAGAGAAATGACATGCTCAAAATC 1237
QY 1172 TGCAGATGACCGGCTGAGATGATCTTCTGCGCTGCGCAAGAGAGCTTAAGCTTC 1231
DB 1238 TCCAGATGACCGGCTGAGATGATCTTCTGCGCTGCGCAAGAGAGCTTAAGCTTC 1297
QY 1232 CGAAAGCAATCTCAAGAAAGAGTCTTCAAGTCTGCTCAAGAGAGTACAGAGAGAGCTTC 1291
DB 1298 CAAGAGCAATCTTCAAGAAAGAGTCTTCAAGTCTGCTCAAGAGAGTACAGAGAGAGCTTC 1357
QY 1292 AGGAATCTGAGCGGCTGCTGATCTCCAGAGAGAGCTTCTCTGCTGATCTCTGCTCC 1351
DB 1358 CGAGCTCAGCCCAATCTCTGCGAGCGCCAGGCGAGCTGCC-----CCCTGCTCC 1408
QY 1352 CAAGAAAGCAATCTTAAAGATCTGCAAGAGTGAATCTGATTAATCTCTCTCCAG 1411
DB 1409 CCAAGAGAGCAATCTTCAAGAAAGCTCCAGAGAGAGAGTGTGATCTCTCTCCG 1468
QY 1412 AGCCAGAGAGTCTGGGAACTTTAGAGCGCAGATGATGATGATGATGATGATGATGATG 1471
DB 1469 AGCCAGAGATCTGGGAACTTTAGAGCGCAGATGATGATGATGATGATGATGATGATGATG 1528
QY 1472 TGAAGCAAGTCTCAAGAGCTTCAAGAGCTCTCTCTCAAGAGAGAGAGAGAGAGAGAGAG 1531
DB 1529 AGAGAGAGAGAGCTCCGCAAGCTTCAAGAGCTCTCTCTCAAGAGAGAGAGAGAGAGAGAG 1588
QY 1532 TCAATGAGAGATCTCCGCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591
DB 1589 TCAATGAGAGATCTCCGCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1648
QY 1592 TGAACCACTGAGCTCTCTCCATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1651
DB 1649 TGAATGAGATCTGAG 1708
QY 1652 GTGAGAGAGAGATCTGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1711
DB 1709 GCGAGAGAGAGATCTGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1768
QY 1712 CGAAAG 1771
DB 1769 CAG 1828
QY 1772 CCTCAGAG-----GTCTGAAGCAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
DB 1829 CCTCAG 1888
QY 1820 GCTTTTCTGAG 1879
DB 1889 GCTTTTCTGAG 1948
QY 1880 GCTCAAGAGTCAAGC 1893
DB 1949 GCTCAAGAGTCAAGC 1962

RESULT 13
CQ714327 1887 bp DNA linear PAT 03-FEB-2004
DEFINITION
SEQUENCE 261 from Patent WO20068579.
ACCESSION
VERSION
CQ714327.1 GI:42275184
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE
1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	Kits, such as nucleic acid arrays, comprising a majority of

```

JOURNAL      Patent: WO 02068579-A 261 06-SEP-2002
FEATURES     PE Corporation (NY) (US)
              location/Qualifiers
              1. .1887
source

```

ORIGIN

Query Match	71.7%;	Score 1357;	DB 6;	length 1687;
Best Local Similarity	84.1%;	Pred. No. 0;		
Matches 1563; Conservative	0;	Mismatches 275;	Indels 21;	Gaps 2

OY	4	CTCTCGCCCTTGCCTCTCGAGAGAGCCCGGCGCTGAGCGAGCGGCTCATCAAGTCTGCTA	106
Db	35	CCACTCCCTCCGGCCCGGAGACTTAGCCCGGCGCTGGCGGAAAGGGCTGATCAATCGGCCCA	94
OY	107	AACCTCTGATGAAAGAGAGCGGCTGTAACCTGGACCATCACAAAACAACCTCGGGAC	166
Db	95	AGCCCTCATATGAAGAAGAGCGGCTGAACCGGACCAACAACAACCTCGGGAC	154
OY	167	GCTACGAGTTCCTGGAGACGCTGGGCAAGGGCACTTACGGGAAAGTGAAAGGCAACGAG	226
Db	155	GCTACGAGTTCCTGGAGACCTCGGGCAAAAGCACTTACGGGAAAGTGAAAGGCGGG	214
OY	227	AGAGCTCGGGGCGCTCGTGGCCATCAAGTCCATCAGAAAGAACAATCAAAAGATAGC	286
Db	215	AGAGCTCGGGGCGCTCGTGGCCATCAAGTCCATCCGGAAGACAAATTAAGAATAGC	274
OY	287	AGGATCTGTGCACATACGGAAGGAGATTGAGATCATGTCTTCACTCAACCCGCCA	346
Db	275	AAGATCTGATGCATACGGAAGGAGATTGAGATCATGTCACTCAACCCCTACA	334
OY	347	TCATTGCCATCATGAAGTGTGGAGAAATGACGACGAAGTTGTGATTTGATGAGATAG	406
Db	335	TCATTGCCATCATGAAGTGTGGAGAACGACGAAAGTCGTGATGTCATGAGATAG	394
OY	407	CCAGCGAGGCGCATCTGTATGATTAATACATGAGAGGGCCACGGCTGAGGACCGGGACG	466
Db	395	CCAGCGGCGGCGCACTTATATGACTATACAGCGAGGCGACGCTCAGTGAACGCGAAG	454
OY	467	CCAGGCAATTTCTTCCGACAGATCGTGTCTGGCCCTGCACATCTGCAACAGAGATCG	526
Db	455	CTAGGCAATTTCTTCCGCGAGATGTCTCTGCCGTGCATATTCGCATCAGAACAGAGTTG	514
OY	527	TTACCGAGATCTCAAGCTGGAAMAACATCTTTAGATGCAATGGAATGMAATCAAGATTG	586
Db	515	TCCACCGAGATCTCAAGCTGGAAGAAATCTCTTGGATGCAATGGAAATATCAAGATTG	574
OY	587	CTGACTTTGGCCTCTTCAACTGTACCAACAAGGCAAGTTCTCTCAGACGTTCTGTGGGA	646
Db	575	CTGACTTTGGCCTCTTCAACTGTACCACTTCACTCAAGGCAAGTTCTGTGGGA	634
OY	647	GCCCTCTCAAGCCCTCGCTGATATGTCAACGGGAAGCCCTATGTGGGCGCCAGAGTTG	706
Db	635	GCCCTCTCATGCTCTCGCCAGAGATGTCAATGGGAAGCCCTTCAACAGGCCAGAGGTGG	694
OY	707	ACAGCTGATCTTGGGCGTTCCTCTGATACCTCTGGTCAATGGCAACATGCCCTTTGACG	766
Db	695	ACAGCTGATCTTGGGCTGTCTCTCTCAATCTCTGGTCAATGGCAACATGCCCTTTGATG	754
OY	767	GCGAGATCATAAACACTGTTGAGCAAAATCACTAACCGGCGTTTACCGTGAACCGGCCA	826
Db	755	GCGATGACCATTAAGATCTTAGTAAACAAGATCAGCAACGGGCGCTTACCGGGAACCACTA	814
OY	827	AGCCGTCCGATGCTGTGGCCTGATCCGGTGGCTGTTAAATGTGTGAACCCCAACCGTGGG	886

[illegible]

RESULT 14
CQ782778

LOCUS C0782778 3395 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2918 from Patent EP1396543.
ACCESSION C0782778
VERSION C0782778.1 GI:45502721
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Oka, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2918 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source 1. .3395
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
118. .2004
/note="unnamed protein product"
/codon_start=1
/protein_id="CAFA6008.1"
/db_xref="GI:45502722"
/translation="MESLVFARRSGTPPSAELARPLAELIKSPKLNKQAVKHH
HKNLNRHYEFLETKGTGYKVKARBSGRVLAJSIRKDIKOEODIMHRIE
HNSLNHPHIIAHEVENSISKIVMEYVSRDLVIERGOLSEREDMRPFIIV
SAVYCHONRVHRDLKLENIILADANGIKIADPGSLNHQKFLIOTGSPILVAP
EIVNGKPYGPEVDSWSIGVLIYLVHGTMPDGHDKILVNOISGAYAREPKSDA
CGILIRLVNPNRTRATLEDVASHMWNKGYATRVBQEAPEHGHGSDASAMAD
WLRSSRPLENGAKVCSFFKQHPAGSGSTPGLEROLHSIKSRKNDMAQSLHSTA
DIDAHKPGKSNLKLPGKILKVKYSAGAEVQEDPRLSPIDPAGQALIPKKGILK
KPKORSGYSSPESGESLIDAGVFPVSGDEKQKPPQASGLILHKGILNLEKP
SOTALELAAPTFGSLDELAPRLPARASPSGAVSEDSITSESDLDLPERLNEP
PLRGCVSNLTGLEBPBSPGSCILRRHRQDPLGSCPSLTDQGVATRYQALNVC
SKLT"

ORIGIN
Query Match 71.7%; Score 1357; DB 6; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
47 CCTCCGCTGAGCTGAGAGCGCCGCGCTGCGGAGCGGCTTCATCAAGTCCTTA 106
152 CCACTCCCTCGCGCGAGAGCTAGCCGCGCTGCGGAGCGGCTTCATCAAGTCCTTA 211
107 AACCTCGATGAAGAGAGCGAGCGGTGAAGCGGACCATCAACAACACTCGGACAC 166
212 AGCCCTTAATGAAGAGAGCGAGCGGTGAAGCGGACCATCAACAACACTCGGACAC 271
167 GCTACGAGTTCCTGAGAGCGCTGGCAAGGCACTTACGGAAGTGAAGAGCGAG 226
272 GCTACGAGTTCCTGAGAGCGCTGGCAAGGCACTTACGGAAGTGAAGAGCGAG 331
227 AGAGCTCGGGCGCTGCTGGTGCATCAAGTCCATCAAGAAACAAATCAAGATGAG 286
332 AGAGCTCGGGCGCTGCTGGTGCATCAAGTCCATCAAGAAACAAATCAAGATGAG 391
287 AGAGTTCGTCGACATACGAGAGGAGATTCATGATTCATCTCAACACCCCA 346
392 AAGATTCGATGACATACGAGAGGAGATTCATGATTCATCTCAACACCCCA 451
347 TCATTGCCATCCATGAAGTGTTCGAGAAATGACAGCAAGATTCATGATTCATG 406
452 TCATTGCCATCCATGAAGTGTTCGAGAAACGACAGCAAGATTCATGATTCATG 511
407 CCAAGCGAGCGCATCTGTATGATTCATCAAGTACGAGCGGCTGATGAGCGGAG 466
512 CCAAGCGAGCGCATCTGTATGATTCATCAAGTACGAGCGGCTGATGAGCGGAG 571
467 CCAAGCATTTCTTCGACAGATGATGCTGCTGACATTCGACCAAGACGGAATG 526

Db 572 CTAGGCAATTTCTTCGAGAGATGCTCTGCGGTGACATTCGACATCGAAGAGTTG 631
527 TTCAACGAGATCTCAAGCTGGAACACATCTTCATAGATCCATGGAACATCAAGATTG 586
632 TCACCGAGATCTCAAGCTGGAACATCTTCATAGATCCATGGAACATCAAGATTG 691
587 CTGACTTTGGCTCTTCAACCTGTACCAAGGCAAGTTCTCTCAAGCGTTCTGTGGGA 646
692 CTGACTTTGGCTCTTCAACCTGTACCAAGGCAAGTTCTCTCAAGCGTTCTGTGGGA 751
647 GCCCTCTACGCTCGCTGAGATGATGACAGGGAAGCCATATGAGGCCCAAGAGTGG 706
752 GCCCTCTACGCTCGCTGAGATGATGACAGGGAAGCCATATGAGGCCCAAGAGTGG 811
707 ACAGCTGTCTCTGAGCGTCTCTCTGTACATCTGCTGTGACCAATGCTCTTGGAG 766
812 ACAGCTGTCTCTGAGCGTCTCTCTGTACATCTGCTGTGACCAATGCTCTTGGAG 871
767 GCGAGATCATTAACAATGCTGAGCAATCATGACGAGGCTTACCGTGAAGCCGCA 826
872 GCGATGACATTAAGATCTGAGAAACAGATCAGCAAGGAGCTTACCGGAGCCACTA 931
827 AGCCGTCGATGCTGTGAGCTGATTCGCTGCTGTTAATGATGAAACCCACCGTGG 886
932 AACCTCTATGCTGTGAGCTGATTCGCTGCTGTTAATGATGAAACCCACCGTGG 991
887 CCACATGAGAGATGATGACATGATTCGCTGCTGTTAATGATGAAACCCACCGTGG 946
992 CCACTCTGAGATGATGACATGATTCGCTGCTGTTAATGATGAAACCCACCGTGG 1051
947 GCGAAGAGAGCTCTGAGAGGCTGAGGACCTTATGATGAACTTTGAGCGGCTTCA 1006
1052 GAGAGAGAGAGCTCTGAGAGGCTGAGGACCTTATGATGAACTTTGAGCGGCTTCA 1111
1007 TGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
1112 TGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171
1067 GCTTCTTCAAGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126
1172 GCTTCTTCAAGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1231
1127 ATTCTTAAAGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1186
1232 ATTCTTAAAGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1291
1187 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
1292 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1351
1247 AGAAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1306
1352 AGAAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1411
1412 TCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1462
1307 TCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1366
1412 TCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1462
1367 TTAAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1426
1463 TTAAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1522
1427 GGAAGCTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1486
1523 GGAAGCTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1582
1487 CAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1546
1583 CAGAGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1642
1547 CCGGAGGCTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1606

Db	1643	CCAGAGAGCGTTGGAGCTGGGGGCCCCAACACACTTGCGGCTCCCTGGATGAATCGGGC	1702
Qy	1607	CCTCCCATCTCTGACAGCCCGGCGCCAGCGCCCTCTAGGGGCTGTGATGATGAAGACATTC	1666
Db	1703	CACCTCGCCCCCTGGCCCGGGCAGCCCAACCTCTAGGGGCTGTGAGCGAGGACAGCATCC	1762
Qy	1667	TGTCCTCCGAGTCTCTTTGACCAATTGACTTGGCCTGAACGTCTTCCCGMAACCCCATCTGA	1726
Db	1763	TGTCCTCTGAGTCTTTTATGACAGCGTGAATTGGCTTGGCTGAACGGGTCCAGAGCCCCCATGTC	1822
Qy	1727	GGGCGTGTGTCTCTGTGACAACTGAGGGGGCTTGAGCAGCTTCCCTCAGAG-----	1780
Db	1823	GGGGCTGTGTCTCTGTGACAACTGACGGGGCTTGAGGAGACCCCCCTCAGAGGGCCCTG	1882
Qy	1781	-----GTCTGAAGCGATGTGTGGCAGGAATCTTTGGGGGATAGCTGCTTTCTCTGACAG	1834
Db	1883	GAGCGTGGCTAGAGGGCGTGGGGGACGAGATCTCTTTGGGGGACAGCTGCTTTTCCCTGACAG	1942
Qy	1835	ACTGCCAAGAGGTGACTGACGCTTACAGACAAAGCCTTAGAAATGTGCTCAAAAGCTCAGC	1893
Db	1943	ACTGCCAAGAGGTGACAGCACTTACCGACAGGCACTAGAGGATGTGCTCAAAAGCTCACCC	2001

RESULT 15			
BD127309			
LOCUS	3395 bp	DNA	linear
DEFINITION	BD127309 Primer for synthesizing full-length cDNA and use thereof.		
			PAT 18-SEP-2002

ACCESSION	BD127309
VERSION	BD127309.1
KEYWORDS	GI:23222254
SOURCE	JP 2002017375-A/2740.
ORGANISM	<i>Homo sapiens (human)</i>
	Home sapiens

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3395)
Ota.T., Nishikawa.T., Isogai.T., Hayashi.K., Ishii.S., Kawai.Y.,
AUTHORS
REFERENCE

TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2740 22-JAN-2002;
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.

COMMENT	HELIOS RESEARCH INSTITUTE
OS	Homo sapiens (human)
PN	JP 2002017375-A/2740
PD	22-JAN-2002
PF	07-JUN-2000 JP 20000251172

PI TOSHIO OTO, TETSUO NISHIKAWA, TAKAO ISOGL, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,

PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10,

PRATIMES	Location/Qualifiers	Key
PC	C12P1/02, C12Q1/68//C12P1/08, G6PF1/30, C12N15/00, C12N5/00	CC
Primer	for synthesizing full-length cDNA and use thereof	PH
	Location/Qualifiers	
FT	CDS (118) . (2001) .	

```

SOURCE
1. 3395
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

ORIGIN	
Query Match:	71.7%; Score 1357; DB 6; Length 3395;
Best Local Similarity	84.1%; Pred No. 0;
Nearest Genbank	AF010363.1; 99.9% identity; 3377 bp; 3395 bp

	Matches	1563;	Conservative	0;	Mismatches	2/57;	Indels	21;	Gaps	2;
Qy	47	CCTCGCCCTGGCCTTCGGAGAGCGGCCGCGCTGGCGGACGCGCTCATCAAGTGCCTTA	106							
Dp	152	CCATCTCCCTGGGCGCCGAGGTAGCGCCGCGCTGGCGGAGAGGGCTGATCAAGTGCCTCA	211							

107 AACCTGATGAGAAGCAGCGGTGAAGCGCACCATCAACAACAACTGCGGACC 166

Db	212	AGCCCTTAATGAAGAAGAGCGGTGAAGGGGACACACACAAGCACACCTGGGCAACC	271
Qy	167	GCTACGAGTTCTCTGGAACGCTGGGCAAGGGCACCCTACGGGAAGTGAAGAAGCACAG	226
Db	272	GCTACGAGTTCTTGGAAACCTCTGGGCAAGGCACCTACGGGAAGTGAAGAAGCGGGG	331
Qy	227	AGACCTGGGGGCTCTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAAAGATAGC	286
Db	332	AGAGCTGGGGGCGCTGGTGGCCATCAAGTCAATCGGAAGACAAATCAAAAGATAGC	391
Qy	287	AGAGTCTGCTGCACATACGAGGAGAGATTGAGATCATGTCTTCACTCAACACCCCCACA	346
Db	392	AAGATCTGATTCACATACGAGGAGATTGAGATCAATGATATCACTCAACACCCCTACA	451
Qy	347	TCATTGCCATCATGAAGTGTTTAGAGATAGACGACAAGATTGATGTCATGAGATATG	406
Db	452	TCATTGCCATCATGAAGTGTTTAGAGACAGCAGCACAAGATCTGATGTCATGAGATATG	511
Qy	407	CCAGCCGAGCGCATCTGTATGATTACATCAGTGAAGCGGCCACGGCTGAGTGAAGCGGACG	466

[illegible][illegible]

	Qy	Db
587	CTGACTTTGGCCTCTCCAACTGTACCAAAAGGCAAGTTCTTCAGACGTTCTGTGGGA	646
652	CTGACTTCGGCCTCTCCAACTCTACCATCAAGGCAAGTTCTTCGACACATTTCTGTGGGA	751

QY 647 GCCCTTCTACGCTCCGTAGATAGTCACCGSAAAGCCTATGTGGGCCCAAGGTGG 706
|||||
|||||
Db 752 GCCCCCTATAGCCTCGGCAGAGAATTGTCAATGGAAGCCTTACAAGGCCCAAGGTGG 811
|||||
|||||

```
Oy      707   ACAGCTGGTCTCTGCGGCATTTCTCCTGAACATTCCGTGTGTCATGSGAACCATGCCCTTTAAGC    766
       |||||
Db      812   ACAAGTGACTCCCTGGGAGTTTTCTCTTAACAATCTGTGTGATGAGCAACCATGGCCTTTAAATG    871
```

OY 767 GGCAGATCATAAACAACCTGTTGAAGAATAACGGGGCTTACCCTGAGCGGCCA 826
 |||||
Db 872 GGCAATGACCAAAGTCCTTAGTAACAGATCAGCAACGGGGCTTACCGGGAGCCACTTA 931

887 CCAACCTGAGAGTATGATGACCACTGATATGAGAGGAGCAACTGGGGGTTATACCAACCGGAGAGCG 945
07 AGCGGTCCGAGACCTGTGTGAGCCCTGATACCGGATGAGCTGTATATGATGATACACCCACCCCGTGTGAG 886
827
Db 932 AACCTCTGATGACCTGTGAGCCTGATACCGGATGAGCTGTGTGATGATGATACCCCAACCCGCGCGGG 991

[illegible]

1052 GAGAGCAGAGGCTCCGAGTAGGGGTGGGCACTCTGGCAGTACTCTGCCCGCGCTTCA 1111

Db	1112	TGGCTGACTGCTCCGGCTTCTTCCGCCCCCTCTGGAAATGGGCGCAAGGTGTGCA	1171
Qy	1067	GCTTCTTCACACACACGTCGCCGAGGTGGACACATGTAACTGGCTGGAGCGGCAAC	1126

db 1172 GCTTCTTCAMACAGATGACCTGTGTGGGGAMACACACCCCTGTGGCTGGAGGCCAGC 1231

QY 1127 ATTCTCTTAAGAGTCCCGAAGGAGGATATGATGGCTCAAAATCTGCAGAGTACCCGG 1186

Db 1232 ATTGCGTCAMAGATCCGCGAMGAGATGCAATGGCCCGCTCTCCACAGTACACGG 1281

Qy 1187 CTGAGGATACCTCTTTCGCGCCCTGGCAGAGACGCTTAAAGCTTCCAAAAGCATTTCTCA 1246

Db 1292 CTGATGACACTGCGCATCGCCCTGGCAAGACAACCTCAGCTGCCAAGGGCATTTCTCA 1351
 QY 1247 AGAAAAAGTCTTACTCTGTCAGGGAGGTATCAGAGAGACCTTCAGAGAACTCAGACCGG 1306
 Db 1352 AGAAGAGGTGTGAGCTTGCAGAAAGGGGTACAGAGAGACCTCCGAGCTCAGGCCAA 1411
 QY 1307 TGGCTGATACTCCAGGGCAGCCTGTGCTGTATCCCTGATCCCAAGGAAGGCATCC 1366
 Db 1412 TCCCTCGAGCCCGAGGGCAGGCTGCCC-----CCTGTCTCCCAAGAGGGCATTC 1462
 QY 1367 TTAAGAAGTCTCGACAGCCGTGAATCTGTATCTCTCTCCAGAGCCCGAGAGTCTG 1426
 Db 1463 TCAGAAGAGCCCGCAGCCGAGAGCTGGCTACTACTCTCTCCGAGCCCGAGTATCTG 1522
 QY 1427 GGGAACTCTTAAAGCCGAGTATGTGTTGTGAGTGGGAGCCCGTGGAGCAGAAGTCTC 1486
 Db 1523 GGGAGCTCTTGGACGACGAGCGATGTGTGTGTGAGTGGGATCCCAAGAGCAGAAGCTTC 1582
 QY 1487 CACAGGCTTCAGGGCTCTCTCTCCACCGCAAGGGCATTCGAAACTCAATGGCAAGTCT 1546
 Db 1583 CGAAGCTTCAGGGCTGCTGCTCTCATCGCAAGGCAATCTCAATGCAAGTCT 1642
 QY 1547 CCGGACAGCCTTAAAGGCACTACCCCTAGCACTTTGCTCCTGACCAACTGAGCTT 1606
 Db 1643 CCGAGACAGCCTTGGAGCTGCGGCCCGACCACTTCGGCTCCTGGATGAATGCGCC 1702
 QY 1607 CCTCCATCTGACAGCCCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAAGACAGCATCC 1666
 Db 1703 CACTTGCCCCCTGGGCCCGGCGCAGCCCTCAGGGGCTGTGAGCGAGACAGCATCC 1762
 QY 1667 TGTCTCCGAGTCTTGAACCAATTGAACTTGCTGAAAGTCTTCCCGAAACCCCACTGA 1726
 Db 1763 TGTCTCTGAGTCTTTGACCACTGAACTTGCTGAAAGCTTCCAGAGCCCCCACTGC 1822
 QY 1727 GGGGCTGTGTGTGTGAAACAACCTGAGGGGCTTGAAGCAGCTCCCTCAGAAG----- 1780
 Db 1823 GGGGCTGTGTGTGTGAAACAACCTGAGGGGCTTGAAGAGCCCCCTCAGAGGGCCCTG 1882
 QY 1781 -----GTCTGAAGCGATGTGGCAGGAATCCTTGGGGGATAGCTGCTTTCTTGAAG 1834
 Db 1883 GAAGCTGCTGAGGCGCTGGCGGCGAGATCTTGGGGGACAGCTGCTTTCCCTGACAG 1942
 QY 1835 ACTGCCAAGAGGTGACTGACGCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
 Db 1943 ACTGCCAAGAGGTGACAGGCACTTACGACAGGCACTGAGGGTGTGCTCAAAAGCTCAC 2001

Search completed: May 15, 2005, 22:46:06
 Job time : 8121.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 16:05:16 / Search time 9307.58 Seconds

(without alignments)
11868.013 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902
Sequence: 1 cactagtgcacccaagat.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC
Gapop 10%0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hsc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2865.8	98.8	2899	3 AK004737	AK004737 Mus muscu
2	2865.8	98.8	3098	3 AK033672	AK033672 Mus muscu
3	2804.4	96.6	2869	3 AK034082	AK034082 Mus muscu
4	1776	61.2	1896	6 AV399041	AV399041 Mus muscu
5	1374.2	47.4	3443	6 HSM801859	AV399039 Homo sapi
6	1264	43.6	1887	9 AV399040	AV399040 Homo sapi
7	951.4	32.8	1887	9 AV399040	AV399040 Homo sapi
8	934.4	32.2	7004	3 CR749209	CR749209 Homo sapi
9	721	24.8	746	4 B1653092	B1653092 60330616
10	720	24.8	795	6 CA319312	CA319312 UI-M-FW0-
11	714	24.6	727	7 CF729223	CF729223 UI-M-HD0-
12	700.2	24.1	992	4 BG143828	BG143828 602815231
13	688	23.7	705	7 CF726196	CF726196 UI-M-GZ0-
14	683.6	23.6	726	6 CB248251	CB248251 UI-M-FD0-
15	672	23.2	672	7 CO424322	CO424322 UI-M-HD0-
16	667.4	23.0	716	7 CO041415	CO041415 UI-M-FW0-
17	663	22.8	986	7 CF584809	CF584809 AGENCOURT
18	660.8	22.6	816	4 B1143828	B1143828 602907107
19	655.4	22.6	704	7 CO041178	CO041178 UI-M-FW0-
20	654	22.5	893	2 BF785290	BF785290 602108490
21	653.2	22.5	713	2 BB625283	BB625283 BB625283
22	652.4	22.5	655	7 CK781308	CK781308 UI-M-G10-
23	651.6	22.5	805	4 BG174288	BG174288 602334589
24	646	22.3	646	4 BG694881	BG694881 NISC_1V09

25	645.8	22.3	948	2 BF789245	BF789245 602105144
26	639.4	22.0	722	4 B1143493	B1143493 602907673
27	639	22.0	669	2 BB078728	BB078728 BB078728
28	627.6	21.6	782	7 CN261003	CN261003 170004247
29	626	21.6	939	2 BE912458	BE912458 601666074
30	624.6	21.5	1069	5 BM927376	BM927376 AGENCOURT
31	615.6	21.2	1022	5 BQ062868	BQ062868 AGENCOURT
32	608	21.0	608	7 CK625537	CK625537 mJ10904.Y
33	600.8	20.7	616	4 B1647932	B1647932 UI-M-FD0-
34	599.4	20.7	660	7 CO043728	CO043728 UI-M-FD0-
35	596.2	20.5	1005	5 BQ060729	BQ060729 AGENCOURT
36	596.2	20.5	1029	5 BQ059508	BQ059508 AGENCOURT
37	596	20.5	674	5 BQ445704	BQ445704 UI-M-ERO-
38	593	20.4	38	6 BG806175	BG806175 2021-17 M
39	584.2	20.4	673	5 BQ746210	BQ746210 UI-M-ERO-
40	584.2	20.1	1012	5 BQ064231	BQ064231 AGENCOURT
41	582.8	20.1	818	7 CK483904	CK483904 AGENCOURT
42	562	19.4	1042	5 BM909401	BM909401 AGENCOURT
43	558.8	19.3	667	4 B1648260	B1648260 603378202
44	558.6	19.2	865	4 BG172224	BG172224 602333366
45	556	19.2	556	6 BY704937	BY704937 BY704937

ALIGNMENTS

RESULT 1	AK004737	2899 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK004737				
DEFINITION	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013B22 product:weakly similar to PROABSE SERINE/THREONINE-PROTEIN KINASE KIA00537 (BC 2.7.1.-) [Homo sapiens], full insert sequence.				
ACCESSION	AK004737	GI:26334437			
VERSION	AK004737.2				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, T., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
REFERENCE	5				

match 154

AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573. (2002)
REFERENCE	6 (bases 1 to 2899)
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,

QY 991 TGAACCCACCCGTCGGGCACTGAGAGATGATAGCCAGTCATTGTGGTCAACTGGG 1050
Db 962 TGAACCCACCCGTCGGGCACTGAGAGATGATAGCCAGTCATTGTGGTCAACTGGG 1021
QY 1051 GTTACACCAACCGAGTGGGGAACAAGAGCCCTGCTGAGGAGTGGGCACTTGTGTG 1110
Db 1022 GTTACACCAACCGAGTGGGGAACAAGAGCCCTGCTGAGGAGTGGGCACTTGTGTG 1081
QY 1111 ACTTGGCCGGGCTCCATGGCGGAGTGTATACGTCGCTCTCGGCCCCCTCTGAGGA 1170
Db 1082 ACTTGGCCGGGCTCCATGGCGGAGTGTATACGTCGCTCTCGGCCCCCTCTGAGGA 1141
QY 1171 ATGAGCCAAAGTGTGAGCTTCTTCAAGCAGACGTCGCGGAGTGGAGCACTGTAC 1230
Db 1142 ATGAGCCAAAGTGTGAGCTTCTTCAAGCAGACGTCGCGGAGTGGAGCACTGTAC 1201
QY 1231 CTGGGCTGAGCGGCAACTTCTTTAAGAGTCCGAAAGAGATGACATGGCTCAA 1290
Db 1202 CTGGGCTGAGCGGCAACTTCTTTAAGAGTCCGAAAGAGATGACATGGCTCAA 1261
QY 1291 ATGCGAAGGTGACCGGCTGAGGATACCTCTTCTGCGCCCTGGAGAGAGCTTTAGC 1350
Db 1262 ATGCGAAGGTGACCGGCTGAGGATACCTCTTCTGCGCCCTGGAGAGAGCTTTAGC 1321
QY 1351 TTCCGAAAGGCAATTCTCAAGAAAAGTCTCTACCTGTCAGGAGGATACAGAGAGAC 1410
Db 1322 TTCCGAAAGGCAATTCTCAAGAAAAGTCTCTACCTGTCAGGAGGATACAGAGAGAC 1381
QY 1411 CTCAGAACTGACCGGCTGATATCTCCAGGCGAGCTGTCTCTGTATCCCTGC 1470
Db 1382 CTCAGAACTGACCGGCTGATATCTCCAGGCGAGCTGTCTCTGTATCCCTGC 1441
QY 1471 TCCCAAGAAAAGATCTTAAAGATCTGACAGCGGATCTGTGTTACTACTCTC 1530
Db 1442 TCCCAAGAAAAGATCTTAAAGATCTGACAGCGGATCTGTGTTACTACTCTC 1501
QY 1531 CAGAGCCAGAGATCTGGGGAATCTTAAAGCAGCAGTATGTTGTGAGTGGGAGC 1590
Db 1502 CAGAGCCAGAGATCTGGGGAATCTTAAAGCAGCAGTATGTTGTGAGTGGGAGC 1561
QY 1591 CCGTGAGCAGAACTCTCAAGGCTTCAAGGCTCTCTCTCACCGAGAGGCAATTCTCA 1650
Db 1562 CCGTGAGCAGAACTCTCAAGGCTTCAAGGCTCTCTCTCACCGAGAGGCAATTCTCA 1621
QY 1651 AACTCAATGCAATGTTCTCCGCAAGCCTTAAAGAGCACTACCCCTTGAACCTTTGCT 1710
Db 1622 AACTCAATGCAATGTTCTCCGCAAGCCTTAAAGAGCACTACCCCTTGAACCTTTGCT 1681
QY 1711 CCGTGAGCAGAACTCTCCGCAAGCCTTCAAGGCTCTCTCTCACCGAGAGGCTG 1770
Db 1682 CCGTGAGCAGAACTCTCCGCAAGCCTTCAAGGCTCTCTCTCACCGAGAGGCTG 1741
QY 1771 TGAAGTGAAGCAGATCTGTCTCTCGAGTCTTGAACAATTGACTGTGCTGAAGCTC 1830
Db 1742 TGAAGTGAAGCAGATCTGTCTCTCGAGTCTTGAACAATTGACTGTGCTGAAGCTC 1801
QY 1831 TTCCGAAACCCCACTGAGGAGCTGTGTCTGTGACAACCTGAGGAGGCTTGAAGC 1890
Db 1802 TTCCGAAACCCCACTGAGGAGCTGTGTGTGTGACAACCTGAGGAGGCTTGAAGC 1861
QY 1891 CTCCCTCAAGAGCTGAAGTGTGAGTGGAGGAATCTTGGGGGATGACTGCTTTCTC 1950
Db 1862 CTCCCTCAAGAGCTGTGAAGTGTGAGTGGAGGAATCTTGGGGGATGACTGCTTTCTC 1921
QY 1951 TGAAGAATGAGAGTGAATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2010
Db 1922 TGAAGAATGAGAGTGAATCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1981
QY 2011 TGAAGTGAAG 2070
Db 1982 TGAAGTGAAG 2041

QY 2071 GAACCTGGGTCGATTTCTCCAGTGAATGATGATCAAGAGGCTTACGTCGAGCC 2130
Db 2042 GAACCTGGGTCGATTTCTCCAGTGAATGATGATCAAGAGGCTTACGTCGAGCC 2101
QY 2131 TGAAGTGAAG 2190
Db 2102 TGAAGTGAAG 2161
QY 2191 TGCCGAGTGTATAGTGGGAGTGCCTGAAGAGTCCCTACCTCTTGTGTCATGATGTC 2250
Db 2162 TGCCGAGTGTATAGTGGGAGTGCCTGAAGAGTCCCTACCTCTTGTGTCATGATGTC 2221
QY 2251 ACCATGACATTTCCACCTCTGTCGTCGACCTTCAATAGTTCTGTTCCAT 2310
Db 2222 ACCATGACATTTCCACCTCTGTCGTCGACCTTCAATAGTTCTGTTCCAT 2281
QY 2311 CAACCAACAGAGTTAAGACCTCTGATCTTCCGAGAGTAAATGTGATGATGATGAT 2370
Db 2282 CAACCAACAGAGTTAAGACCTCTGATCTTCCGAGAGTAAATGTGATGATGATGAT 2341
QY 2371 TAGAGAGAAACAGCTCTGATCTTCCATCTCTGTCGTCGATCTCAAGAGAGAGAG 2430
Db 2342 TAGAGAGAAACAGCTCTGATCTTCCATCTCTGTCGTCGATCTCAAGAGAGAGAG 2401
QY 2431 GACTCGAGCCGCTGTTGACTTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2490
Db 2402 GACTCGAGCCGCTGTTGACTTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2461
QY 2491 GATCTCAG 2550
Db 2462 GATCTCAG 2521
QY 2551 TTCTCAG 2610
Db 2522 TTCTCAG 2581
QY 2611 TCTGTTTTTGTGTTTTGTTGTTTTGTTGTTTTGTTGTTTTGTTGTTTTGTTGTTTT 2670
Db 2582 TCTGTTTTTGTGTTTTGTTGTTTTGTTGTTTTGTTGTTTTGTTGTTTTGTTGTTTT 2641
QY 2671 CAATATGTAATGCTGTGTTCTGGGGAATCTCACTGTGCACTGAAGTTATGATGAGA 2730
Db 2642 CAATATGTAATGCTGTGTTCTGGGGAATCTCACTGTGCACTGAAGTTATGATGAGA 2701
QY 2731 GAAGTATTTGGCAATGATGCTCTCTATTAAGAGGAGGAGGAGGAGGAGGAGGAGG 2790
Db 2702 GAAGTATTTGGCAATGATGCTCTCTATTAAGAGGAGGAGGAGGAGGAGGAGGAGG 2761
QY 2791 GTCTTGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2850
Db 2762 GTCTTGAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2821
QY 2851 CTTGATCTTCACTGTGAGCAAAAAA 2899
Db 2822 CTTGATCTTCACTGTGAGCAAAAAA 2870

RESULT 2
AK033672
LOCUS
DEFINITION
MUS MUSCULUS ADULT MALE CECUM CDNA, RIKEN FULL-LENGTH ENRICHED
LIBRARY, CLONE 9130215K18 PRODUCT:WEAKLY SIMILAR TO PROBABLE
SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) [Homo
sapiens], full insert sequence.
ACCESSION
AK033672
VERSION
AK033672.1 GI:26329364
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS 3
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, T., Nishi, K., Kitzunai, T., Tashiro, H., Itoh, M.,
Sumi, M., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matcalik, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE
AUTHORS 5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 12030981
PUBMED 12030981
REFERENCE
AUTHORS 6
Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P.,
Pekuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsumura, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source location/Qualifiers
1. 3098
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:9130215K18"
/db_xref="taxon:10090"
/clone="9130215K18"

CDs
/sex="male"
/tissue_type="cecum"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
93.1988
/note="unnamed protein product; putative
weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
KIAA0537 (EC 2.7.1.-) (Homo sapiens) (SWISSPROT|O60285,
evidence: FASTA, 56.8% ID, 98.6% length, match=1849) "
/codon_start=1
/protein_id="BAC28421.1"
/db_xref="GI:26329365"
/translation="MRSVALTORPSQAPSAASALSSAPRLADGLIKSPRLKKQAV
KRHHKGNLRHREPIETLAKGTGYKAKRSGSLVAIKSRKIDTDEDLHNR
RELEISLNHPHIIHVEFENSSKIVIMETASRGDLVDYISERPERLRARHP
ROIVSLAHYCHONGIVHRDLKLENIILDNAGNIKIADFGISLNYHKGFLOTGSP
YAPPELVKPYGVPDWSLSGLVLIIVHGTMPDGDHKTLLVKQISNGAYRPPK
PSAPGCIIRLILVNPTRRLTLEDVASHMMVMNNGYTTGVEOELBEGHSPDGFGA
SMADWIRSRRLLENGARVCSFPKOHVGGCSVTPLGLEHQSILKSRKENDMAQNIQ
GDPAEDTSSRPKSSILKPKGIIKKKSSITSSGVEQDPELRFPVPTPPQVPVAVSL
PRGILKRSRQRESGYSSPESESELDAVDVFPVSGDVEKSPQASGLLHARKGI
LKNLGRFSRYALEGTPSTFGSLDOLASHPARPSRPGCAVEDSITLSESPDOLD
PERLPTPLRGCVSVNLRGLBQPEPSEGLKRWQESLGDSFCFLTDCQEVTAAYRQAL
GICSKLS"
polyA_signal 3082..3087
/note="putative"
polyA_site 3098
/note="putative"
ORIGIN
polyA_signal 3082..3087
/note="putative"
polyA_site 3098
/note="putative"
Query Match 98.8%; Score 2865.8; DB 3; Length 3098;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2867; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 31 GCGTGTCTGGGATGCGGCTGTGACCTCTGAGCCCGGCTCAGCGCGCTGTACTGCTG 90
DB 1 GGGTGTCTGGGATGCGGCTGTGACCTCTGAGCCCGGCTCAGCGCGCTGTACTGCTG 60
QY 91 CCCGACCTCACTCCACTCGGCTGCTCCCGGACCATGAGTGGTGGCTTACTCCAGGCC 150
DB 61 CCCGACCTCACTCCACTCGGCTGCTCCCGGACCATGAGTGGTGGCTTACTCCAGGCC 120
QY 151 CGAGCCAGGCTCCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 210
DB 121 CGAGCCAGGCTCCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 180
QY 211 GGGTCAATCAAGTGGCTTAACTCTGATGAAGAAGCAGGCGGGAAGCGGCACATCACA 270
DB 181 GGGTCAATCAAGTGGCTTAACTCTGATGAAGAAGCAGGCGGGAAGCGGCACATCACA 240
QY 271 AACACAACTGCGGACCCGCTACGAGTTCCTGAGACGCTGGGCAAGGACCTTACCGGA 330
DB 241 AACACAACTGCGGACCCGCTACGAGTTCCTGAGACGCTGGGCAAGGACCTTACCGGA 300
QY 331 AGGTGAAGAGGACGAGAGAGTCTGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAG 390
DB 301 AGGTGAAGAGGACGAGAGAGTCTGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAG 360
QY 391 ACAAAATCAAGATGAGAGAGATCTGCTGACATACGAGAGGAGATTGAGATGATGCTT 450
DB 361 ACAAAATCAAGATGAGAGAGATCTGCTGACATACGAGAGGAGATTGAGATGATGCTT 420
QY 451 CACTCAACCAACCCCAATCATTCATCCATCAGTAAGTGTGGAATAGAGCAAGATTG 510
DB 421 CACTCAACCAACCCCAATCATTCATCCATCAGTAAGTGTGGAATAGAGCAAGATTG 480
QY 511 TGAATTGTCATGAGATATGCCAGCGGAGCGATCTGATGATTTACATCAGTGAAGCGG 570
DB 481 TGAATTGTCATGAGATATGCCAGCGGAGCGATCTGATGATTTACATCAGTGAAGCGG 540
QY 571 GGGTGAAGTGAAGGAGCGGAGGATTTCTTCCGACAGATCGGTGCGCTCCATCACT 630
DB 541 GGGTGAAGTGAAGGAGCGGAGGATTTCTTCCGACAGATCGGTGCGCTCCATCACT 600

QY 631 GCCACGAAAGGAGTGTTCACCGAGATCTCAAGCTGGAACATCTTTAGATGCCA 690
Db 601 GCCACGAAAGGAGTGTTCACCGAGATCTCAAGCTGGAACATCTTTAGATGCCA 660
QY 691 ATGGAACATCTAATTTGCTGACCTTTGGCCCTCTCAACCTGTACCAAGAGGATGCC 750
Db 661 ATGGAACATCTAATTTGCTGACCTTTGGCCCTCTCAACCTGTACCAAGAGGATGCC 720
QY 751 TCCAGAGCTTCTGAGGAGCCCTCTCAAGCTGTGAGATGATCAAGAGGAGCCCT 810
Db 721 TCCAGAGCTTCTGAGGAGCCCTCTCAAGCTGTGAGATGATCAAGAGGAGCCCT 780
QY 811 ATGTGGCCCAAGAGTGAAGCTGTCTCTGGGCTTCTCTGATCAATCTGTGATG 870
Db 781 ATGTGGCCCAAGAGTGAAGCTGTCTCTGGGCTTCTCTGATCAATCTGTGATG 840
QY 871 GCACATGCTCTTGAAGGAGATCAATAACAATGTGAGCAATCAATAAGGAGG 930
Db 841 GCACATGCTCTTGAAGGAGATCAATAACAATGTGAGCAATCAATAAGGAGG 900
QY 931 CTTAACGAGAGCGCCCAAGCCGATCCGATGCTGAGCTGATCCGAGCTGTAAATG 990
Db 901 CTTAACGAGAGCGCCCAAGCCGATCCGATGCTGAGCTGATCCGAGCTGTAAATG 960
QY 991 TGAACCCCAACCCGTCGAGGCACTGAGAGATGATCAATGTGTGAGTCAATGAG 1050
Db 961 TGAACCCCAACCCGTCGAGGCACTGAGAGATGATCAATGTGTGAGTCAATGAG 1020
QY 1051 GTTACACCAACCGAGTGGGAGCAAGAAAGCTGTGAGGAGGAGGAGGAGGAGT 1110
Db 1021 GTTACACCAACCGAGTGGGAGCAAGAAAGCTGTGAGGAGGAGGAGGAGGAGT 1080
QY 1111 ACTTTGGCCGAGCTCTCAATGGCCGATGATGCTGCTCTCCGCTCCCTCTGAGAG 1170
Db 1081 ACTTTGGCCGAGCTCTCAATGGCCGATGATGCTGCTCTCCGCTCCCTCTGAGAG 1140
QY 1171 ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGTCGCGAGAGTGAAGCACTGAT 1230
Db 1141 ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGTCGCGAGAGTGAAGCACTGAT 1200
QY 1231 CTGGGCTGAGAGGAGCAATTTCTTTAAGAAAGTCCGAAAGAGATGATCATGCTCAA 1290
Db 1201 CTGGGCTGAGAGGAGCAATTTCTTTAAGAAAGTCCGAAAGAGATGATCATGCTCAA 1260
QY 1291 ATCTGCAAGTGAACCCGCTGAGAGTACTCTTTCTCCGCTGGAAGAGCACTTAA 1350
Db 1261 ATCTGCAAGTGAACCCGCTGAGAGTACTCTTTCTCCGCTGGAAGAGCACTTAA 1320
QY 1351 TTCCGAAAGGATTTCTCAAGAAAGTCTCTACCTGCTCAAGGAGGATCAGAGAGC 1410
Db 1321 TTCCGAAAGGATTTCTCAAGAAAGTCTCTACCTGCTCAAGGAGGATCAGAGAGC 1380
QY 1411 CTCAGAACTCAGACCGGTGCTGATATCTCAAGGAGCAGCTGCTGCTGATCCTGC 1470
Db 1381 CTCAGAACTCAGACCGGTGCTGATATCTCAAGGAGCAGCTGCTGCTGATCCTGC 1440
QY 1471 TCCCAAGAAAGGATCTTTAAGAAAGTCTGACAGCTGATCTGTGATCTACTCTCTC 1530
Db 1441 TCCCAAGAAAGGATCTTTAAGAAAGTCTGACAGCTGATCTGTGATCTACTCTCTC 1500
QY 1531 CAGAGCCAGAGTCTGAGGAGCTTTAGACGCGAGATGATGTTGTTGAGTGGAGCC 1590
Db 1501 CAGAGCCAGAGTCTGAGGAGCTTTAGACGCGAGATGATGTTGTTGAGTGGAGCC 1560
QY 1591 CCGTGAAGCAGAACTTCTCAGAGCTTCAAGGCTCTCTCTCACCAGCAAGGAGATTTCA 1650
Db 1561 CCGTGAAGCAGAACTTCTCAGAGCTTCAAGGCTCTCTCTCACCAGCAAGGAGATTTCA 1620
QY 1651 AACTCAATGCAATTTCTCCGCAAGCTTGAAGAGCACTAACCTTGAAGCCTTTGGCT 1710
Db 1621 AACTCAATGCAATTTCTCCGCAAGCTTGAAGAGCACTAACCTTGAAGCCTTTGGCT 1680

QY 1711 CCTTGAACCAATGAGCTCTCTCCATCTCTGAGCCGCGCCGACCCCTCTCAGAGGCTG 1770
Db 1681 CCTTGAACCAATGAGCTCTCTCCATCTCTGAGCCGCGCCGACCCCTCTCAGAGGCTG 1740
QY 1771 TGAATGAGGAGCAGATCTGCTGCTCGAGTCTTTGACCAATTTGATCTGCTGAAGCT 1830
Db 1741 TGAATGAGGAGCAGATCTGCTGCTCGAGTCTTTGACCAATTTGATCTGCTGAAGCT 1800
QY 1831 TTCCGAAACCCCACTGAGGAGCTGTGTCTGTGACAACCTGAGGAGGCTTGAAGCAGC 1890
Db 1801 TTCCGAAACCCCACTGAGGAGCTGTGTCTGTGACAACCTGAGGAGGCTTGAAGCAGC 1860
QY 1891 CTCCCTGAAGAGTCTGAAGGATGTGTGAGAGCAATCTTTGGGAGATGCTGCTTTTCTC 1950
Db 1861 CTCCCTGAAGAGTCTGAAGGATGTGTGAGAGCAATCTTTGGGAGATGCTGCTTTTCTC 1920
QY 1951 TGAAGACTGCAAGAGTGAATGACGCTACAGCAAGCCCTGAGGAATGCTCAAGC 2010
Db 1921 TGAAGACTGCAAGAGTGAATGACGCTACAGCAAGCCCTGAGGAATGCTCAAGC 1980
QY 2011 TCAGCTGAGGAGGAGATGTGTGCTTATGAGGATAGGCTCTGAGAGGCTTTGACAG 2070
Db 1981 TCAGCTGAGGAGGAGATGTGTGCTTATGAGGATAGGCTCTGAGAGGCTTTGACAG 2040
QY 2071 GAACCTGTGGTCTGATTTCTTCAATGATGATGATCAATCAAGGCTCTTACGCTGACGC 2130
Db 2041 GAACCTGTGGTCTGATTTCTTCAATGATGATGATCAATCAAGGCTCTTACGCTGACGC 2100
QY 2131 TGAATGAACCTGAAGATGAGAGAAATGCAATTTATGTGAAAGAAATGGAAACCTTTGC 2190
Db 2101 TGAATGAACCTGAAGATGAGAGAAATGCAATTTATGTGAAAGAAATGGAAACCTTTGC 2160
QY 2191 TGCCGAGTGTATAGTGGGAGTGCCTGAAGTGCCTAACCTTTTGCCATGAGTGC 2250
Db 2161 TGCCGAGTGTATAGTGGGAGTGCCTGAAGTGCCTAACCTTTTGCCATGAGTGC 2220
QY 2251 ACCATGACATTTTCCACCCCTGATCTGTGCTGACCTTCAATAGTTTGTGTTCCAT 2310
Db 2221 ACCATGACATTTTCCACCCCTGATCTGTGCTGACCTTCAATAGTTTGTGTTCCAT 2280
QY 2311 CAACCAACGAGGTTAGAACCTTGAATCTGTGAGAGTATGTATGTGATCTGCCATTTT 2370
Db 2281 CAACCAACGAGGTTAGAACCTTGAATCTGTGAGAGTATGTATGTGATCTGCCATTTT 2340
QY 2371 TAGAGAGAAACACCTCTGCTTCCATCTGTGCTGTGATCTCTCAAGACCTGGGAA 2430
Db 2341 TAGAGAGAAACACCTCTGCTTCCATCTGTGCTGTGATCTCTCAAGACCTGGGAA 2400
QY 2431 GACTCGAACCGCTGTGATCTTCAATCTCAAGGAGCAAGATGCCCTGAGACCCATCTTA 2490
Db 2401 GACTCGAACCGCTGTGATCTTCAATCTCAAGGAGCAAGATGCCCTGAGACCCATCTTA 2460
QY 2491 GATTCAGAGCTTGAACCTTGAAGCTGTCTTATGATCCAGATGTGATGATGCTGTG 2550
Db 2461 GATTCAGAGCTTGAACCTTGAAGCTGTCTTATGATCCAGATGTGATGATGCTGTG 2520
QY 2551 TTTCTCAGGCAAGGAGCCAGAAATGAGTGAATTTATTTTGTGATCTCACT 2610
Db 2521 TTTCTCAGGCAAGGAGCCAGAAATGAGTGAATTTATTTTGTGATCTCACT 2580
QY 2611 TCTGTGTTTGTGTTTGTGTTGTTGTTGTTTGTGTTTAAAGTGAATTTGTGCTTT 2670
Db 2581 TCTGTGTTTGTGTTTGTGTTGTTGTTGTTTGTGTTTAAAGTGAATTTGTGCTTT 2640
QY 2671 CAATTAATGATGATCTGTGTTCTGAGGAACTCACTGTGCACTGAAGTTTATGATCAGA 2730
Db 2641 CAATTAATGATGATCTGTGTTCTGAGGAACTCACTGTGCACTGAAGTTTATGATCAGA 2700
QY 2731 GAAGTATTTGGCAATGATGCTCTTATTTCAAGGAGGAGTGGGAGGCTTTTCAATGAT 2790
Db 2701 GAAGTATTTGGCAATGATGCTCTTATTTCAAGGAGGAGTGGGAGGCTTTTCAATGAT 2760
QY 2791 GTCTTGAAGCACTGTCTGATTTGATGCTCAAGTCCCTTCAACCAAGGCTGGCAACCTC 2850

Db 2761 GCTTGTGACATGCTGCTGATTTGAGTCTCCAGTCCCTTACACCCCAAGGCTGGCCACCTC 2820

QY 2851 CCTCATCTTCATCTGCTGCGCCAAAAA 2899

Db 2821 CCTCATCTTCATCTGCTGCGCCAAAAA 2869

RESULT 3
AK034082
LOCUS

DEFINITION
AK034082 2869 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330154N24 product:weakly similar to
PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)
[Homo sapiens], full insert sequence.

ACCESSION
AK034082
VERSION
AK034082.1 GI:26329672
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsumi, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076661

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

FEATURES
source
1. 2869
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9330154N24"
/db_xref="taxon:10090"
/clone="9330154N24"
/sex="male"
/tissue_type="dienecephalon"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
98. 2017
/note="unnamed protein product; putative
weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|O60285,
evidence: FASTV, 56.8%id, 98.6%length, match=1849)"
/codon_start=1
/protein_id="BAC28575.1"
/db_xref="GI:26329673"
/translation="MSGALIDPSPQAPSAASASAPLADGLIKSPYMKKQV
KRHHKNNLHREYLETJAGKYGVKAKRESGRVAIKSTRKDKIOEDLHLR
RELTISSLNHPHIAIHEVGRSLVTVPENSSKIVMEVASRGDLYDISRPLRS
BRDARFPRQIVSALHYCHONGIVHRDGLKENTLIDANGNIKIADFGLSNLYKXGL
QTGCGSLPVSPEIVNGKPYGPEVDSWSLGLVLYLHGTGPFQDGHKTLVQISN
PAGRPKPSDAGLIRMLIMVNPTRATLEDVASHMWVWVGTTGVBGEALREGH
PSDRPRPSADPLRRSSRPILNENKAVCFPQHPGGSSTYPLGRQSLSKRE
NDIAQNLQGPADTSSRPKSLKLPKGLTKKSSSTSSSEVEDPELRLPVDTPQ
PVAVSLLPKGLIKSRORSEGYSPSPSESGELDASHPAPRSFSAVSPQASG
LILHARGILKLNKFSRTALEGTPSTFGSLDQLASHPAPRSFSAVSPQASG
ESFDQDLPERLPEPLRGCVSDNLRGLQEPSEGLKRWQSGSLDSCGV
TAAYRQALGICSLTS"

ORIGIN
Query Match 96.6%; Score 2804.4; DB 3; length 2869;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2839; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 31 GCGTGTCTGGGTCGGGTGCTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTGCTG 90

Db 6 GCGTGTCTGGGTCGGGTGCTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTGCTG 65

QY 91 CCGGACCCACTCCACCGCTCGCGGTCCCGCGACCATGAGTGGTGGCTTACTCCAGGCC 150

Db 66 CCGGACCCACTCCACCTCGCGGTCCCGCGACCATGAGTGGTGGCTTACTCCAGGCC 125

QY 151 CGAGCCAGGCTCCCTCGGCTCGGCTCGGCTCGAGAGAGCGCCCGCTGCGGAGC 210

Db 126 CGAGCCAGGCTCCCTCGGCTCGGCTCGGCTCGAGAGAGCGCCCGCTGCGGAGC 185

QY 211 GCGTCATCAAGTGGCTTAACTCTGTGATGAAGAAGCGCGGTGAAGCGGCACCATACA 270

Db 186 GCGTCATCAAGTGGCTTAACTCTGTGATGAAGAAGCGCGGTGAAGCGGCACCATACA 245

QY 271 AAGACAACCGGGGACCGGTAGAGTCTCGAGAGCGCTGGGCAAGGGGCACTACGGGA 330

Db 246 AAGACAACCGGGGACCGGTAGAGTCTCGAGAGCGCTGGGCAAGGGGCACTACGGGA 305

QY 331 AGGTGAAGAAAGGACGAGAGAGCTCGGAGGCTGTGTGGCCATCAAGTCCATCGAGAAAG 390
 Db 306 AGGTGAAGAAAGGACGAGAGAGCTCGGAGGCTGTGTGGCCATCAAGTCCATCGAGAAAG 365
 QY 391 ACAAATCAAAAGATGACGAGATGCTGTGCATATCGAGAGGAGATTTGAGTCAATGTCTT 450
 Db 366 ACAAATCAAAAGATGACGAGATGCTGTGCATATCGAGAGGAGATTTGAGTCAATGTCTT 425
 QY 451 CACTCAACCAACCCCACTCATCTGCAATCCATGAA-----G 486
 Db 426 CACTCAACCAACCCCACTCATCTGCAATCCATGAAAGTGGAGAGATCAACGCTTAAGTACTG 485
 QY 487 TGTTTGAAGATGACGAGAGATTTGATTTGATCATGAGATATGCAAGCCGAGCATCTGT 546
 Db 486 TGTTTGAAGATGACGAGAGATTTGATTTGATCATGAGATATGCAAGCCGAGCATCTGT 545
 QY 547 ATGATTTATCATGATGAGAGAGGACCGGCTGAGTGAAGCCGAGACGCAAGGCAATTCCTGAC 606
 Db 546 ATGATTTATCATGATGAGAGAGGACCGGCTGAGTGAAGCCGAGACGCAAGGCAATTCCTGAC 605
 QY 607 AGATCGGTGCGCCCTGCACTACCTGCAACGAGAGGAGATGTTCAACGAGATCTCAAGC 666
 Db 606 AGATCGGTGCGCCCTGCACTACCTGCAACGAGAGGAGATGTTCAACGAGATCTCAAGC 665
 QY 667 TGAAGAAATCATCTCTCTAGATGCAATGAGAAATCAAGATTTGCTGACTTTGGCTCTCA 726
 Db 666 TGAAGAAATCATCTCTCTAGATGCAATGAGAAATCAAGATTTGCTGACTTTGGCTCTCA 725
 QY 727 ACTGTATCAAAAGGCAAGTTCTCTCAAGACGTTCTGTGGAGCCCTCTCAAGCCTCTGC 786
 Db 726 ACTGTATCAAAAGGCAAGTTCTCTCAAGACGTTCTGTGGAGCCCTCTCAAGCCTCTGC 785
 QY 787 CTGAGATAGTCAAGGAGAGCCCTATGAGGAGCCAGAGTGAAGAGCTGAGTCTGAGGCG 846
 Db 786 CTGAGATAGTCAAGGAGAGCCCTATGAGGAGCCAGAGTGAAGAGCTGAGTCTGAGGCG 845
 QY 847 TTCTCCGTATCATCTGTGTGATGAGCAACATGACCTTTGACGAGGACGAGATCATTAACAC 906
 Db 846 TTCTCCGTATCATCTGTGTGATGAGCAACATGACCTTTGACGAGGACGAGATCATTAACAC 905
 QY 907 TGTGAAGCAAAATGATGACGAGGCTTACCGTGAAGCCGCAAGCCGCTCGAGTCTGTG 966
 Db 906 TGTGAAGCAAAATGATGACGAGGCTTACCGTGAAGCCGCAAGCCGCTCGAGTCTGTG 965
 QY 967 GCGTGAATCCGAGTGTGTTAATGATGAAACCCCAACCCGCTCGAGGACATGAGAGATGAG 1026
 Db 966 GCGTGAATCCGAGTGTGTTAATGATGAAACCCCAACCCGCTCGAGGACATGAGAGATGAG 1025
 QY 1027 CCAAGTCAATGGTGGTCAACGAGGAGTTTACCAACCGAGTGGGAGAACAGGAGCCCTGC 1086
 Db 1026 CCAAGTCAATGGTGGTCAACGAGGAGTTTACCAACCGAGTGGGAGAACAGGAGCCCTGC 1085
 QY 1087 GTGAGGATGAGCAACCTAGTGTGATCTTTGACCGGAGCTCCATGAGCGAGCTGTGTAAGTC 1146
 Db 1086 GTGAGGATGAGCAACCTAGTGTGATCTTTGACCGGAGCTCCATGAGCGAGCTGTGTAAGTC 1145
 QY 1147 GCTCTCTGCGCCCTCTCTGAGAGATGAGCAAGGTGTGAGCTTTTCAAGAGAGACG 1206
 Db 1146 GCTCTCTGCGCCCTCTCTGAGAGATGAGCAAGGTGTGAGCTTTTCAAGAGAGACG 1205
 QY 1207 TGCCGAGAGGTGAGAGCACTGTACCTGGAGCTGAGCGGCAACATTTCTTAAGAAAGTCC 1266
 Db 1206 TGCCGAGAGGTGAGAGCACTGTACCTGGAGCTGAGCGGCAACATTTCTTAAGAAAGTCC 1265
 QY 1267 GAAAGAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGCTGAGAGATACCTCTCTC 1326
 Db 1266 GAAAGAGAGATGACATGCTCAAAATCTGCAAGGTGACCCGCTGAGAGATACCTCTCTC 1325
 QY 1327 GCGCTGCAAGAGAGAGCTTAAAGCTTCCGAAAGGCAATTCGAAGAAAGTCTCTACT 1386
 Db 1326 GCGCTGCAAGAGAGAGCTTAAAGCTTCCGAAAGGCAATTCGAAGAAAGTCTCTACT 1385

QY 1387 CGTCAAGGAGAGATACAGAGAGAGCCCTCAAGGAATCTCAACCGGAGCTGATATCTCAAGGAC 1446
 Db 1386 CGTCAAGGAGAGATACAGAGAGAGCCCTCAAGGAATCTCAACCGGAGCTGATATCTCAAGGAC 1445
 QY 1447 AGCGTGCCTGCTGTATCCCTGTGCTCCAGAGAAAGGATCTTTAAGAGTCTGACAGC 1506
 Db 1446 AGCGTGCCTGCTGTATCCCTGTGCTCCAGAGAAAGGATCTTTAAGAGTCTGACAGC 1505
 QY 1507 GTGAATCTGTGTTACTACTCTCTCCAGAGCCAGCGAGTGTGGAGAACTTTAGACGCCA 1566
 Db 1506 GTGAATCTGTGTTACTACTCTCTCCAGAGCCAGCGAGTGTGGAGAACTTTAGACGCCA 1565
 QY 1567 GTGATGTGTTGTGAGTGGAGAGCCCGTGAAGAGATCTTCAAGGCTTCAAGGCTTCC 1626
 Db 1566 GTGATGTGTTGTGAGTGGAGAGCCCGTGAAGAGATCTTCAAGGCTTCAAGGCTTCC 1625
 QY 1627 TCCCTCAACCGCAAGGAGATCTCAAAATCAATGAGCAAGTTCTCCGCAAGCCCTTAAG 1686
 Db 1626 TCCCTCAACCGCAAGGAGATCTCAAAATCAATGAGCAAGTTCTCCGCAAGCCCTTAAG 1685
 QY 1687 GCACTTACCCCTAGACCTTTGAGCTCCCTGAGACCACTGAGCTCTCCATCTTCAAGCC 1746
 Db 1686 GCACTTACCCCTAGACCTTTGAGCTCCCTGAGACCACTGAGCTCTCCATCTTCAAGCC 1745
 QY 1747 GAGCCAGAGCCCTCTAGAGGAGCTGTGAGTGAAGAGACAGATCTGTCTCTCGAGTCTTTG 1806
 Db 1746 GAGCCAGAGCCCTCTAGAGGAGCTGTGAGTGAAGAGACAGATCTGTCTCTCGAGTCTTTG 1805
 QY 1807 ACCAATTTGATCTTGTCTGAAAGTCTTCCGAAACCCCACTGAGAGGAGCTGTGTCTGTG 1866
 Db 1806 ACCAATTTGATCTTGTCTGAAAGTCTTCCGAAACCCCACTGAGAGGAGCTGTGTGTGTG 1865
 QY 1867 ACAACCTGAGAGGAGCTTGAAGAGCCCTCTCAAGAGTCTGAAGCGATGTGAGAGAT 1926
 Db 1866 ACAACCTGAGAGGAGCTTGAAGAGCCCTCTCAAGAGTCTGAAGCGATGTGAGAGAT 1925
 QY 1927 CTTTGGAGAGATAGCTGCTTTTCTCTGAAGAGCTGCAAGAGTCTGAAGCGATGTGAGAG 1986
 Db 1926 CTTTGGAGAGATAGCTGCTTTTCTCTGAAGAGCTGCAAGAGTCTGAAGCGATGTGAGAG 1985
 QY 1987 AAGCCCTAGAGATTTGCTCAAAAGTCAAGCTGAGAGAGAGAGATGTGCTCTGATGTGGG 2046
 Db 1986 AAGCCCTAGAGATTTGCTCAAAAGTCAAGCTGAGAGAGAGAGATGTGCTCTGATGTGGG 2045
 QY 2047 TAGGCTGAGAGGAGTTGAGAGAGAACCTGAGGCTGAGATCTTCAAGTGAATGAGATC 2106
 Db 2046 TAGGCTGAGAGGAGTTGAGAGAGAACCTGAGGCTGAGATCTTCAAGTGAATGAGATC 2105
 QY 2107 ATCAAGGAGCTTACGTCTGAGAGCTGAGCTGAACCTGAAAGATGAGAGAAATCGCATTTGAT 2166
 Db 2106 ATCAAGGAGCTTACGTCTGAGAGCTGAGCTGAACCTGAAAGATGAGAGAAATCGCATTTGAT 2165
 QY 2167 GTGAAAGAGATGAGAACCTTGTGCTGCCAGATGTATATGAGGAGGCTGAAAGTGGC 2226
 Db 2166 GTGAAAGAGATGAGAACCTTGTGCTGCCAGATGTATATGAGGAGGCTGAAAGTGGC 2225
 QY 2227 TACCTCTTGTGAGAGAGATGACACCAAGATGATTTCCACCTGTTCTCTGAGCTGAC 2286
 Db 2226 TACCTCTTGTGAGAGAGATGACACCAAGATGATTTCCACCTGTTCTCTGAGCTGAC 2285
 QY 2287 CTTCACATTAAGTTCTGTCTTCCATCAACCAACAGAGGTTAGAACCTGATCTTCCGAGAG 2346
 Db 2286 CTTCACATTAAGTTCTGTCTTCCATCAACCAACAGAGGTTAGAACCTGATCTTCCGAGAG 2345
 QY 2407 CTGTGATCTCAAGAGAGCTGGAGAGAGCTGGAGAGAGCTGGAGAGAGCTTCAATCAAGAGAG 2466
 Db 2406 CTGTGATCTCAAGAGAGCTGGAGAGAGCTGGAGAGAGCTGGAGAGAGCTTCAATCAAGAGAG 2465
 QY 2467 CAGATGCCCTGAGACCCCATCTTAGATCTCAAGAGACTTGAACCTTGAAGCTGTTCTTAGT 2526

Db	2466	CAGATGCCCTGAGACCCCATCTTGAATCTCAGAGACTTGAACCTTGAAGCTGTTCTTAGT	2525		
Qy	2527	ACCCAGATGTGGATGATGATGCTCTGTTTCTCAGGCCAACGAGACCTAGAAATGCTGACTT	2586		
Db	2526	ACCCAGATGTGGATGATGATGCTCTGTTTCTCAGGCCAACGAGACCTAGAAATGCTGACTT	2585		
Qy	2587	ATTATATTTTTTGTGATGATCTCAGCTTCGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTG	2646		
Db	2586	ATTATATTTTTTGTGATGATCTCAGCTTCGTGTTTTTGTGTTTTTGTGTTTTTGTGTTTTTGTG	2645		
Qy	2647	TTTTTAAGTGAATTTTGTGCTTTCAATATATGTAATGTAATGCTGATGTTCTGAGGAACTCCACT	2706		
Db	2646	TTTTTAAGTGAATTTTGTGCTTTCAATATATGTAATGTAATGCTGATGTTCTGAGGAACTCCACT	2705		
Qy	2707	GTGCCACTGAAAGTTATGATACAGAGAAATTTGGCAATGATATGTCCTCTATTCAGAGGG	2766		
Db	2706	GTGCCACTGAAAGTTATGATACAGAGAAATTTGGCAATGATATGTCCTCTATTCAGAGGG	2765		
Qy	2767	GGTGGGGGGCGTTTTCCTCAATATGATGCTTGGAGCACTGCTGATTTGAATCTCCAGTCCCT	2826		
Db	2766	GGTGGGGGGCGTTTTCCTCAATATGATGCTTGGAGCACTGCTGATTTGAATCTCCAGTCCCT	2825		
Qy	2827	TCACACCCCAAGAGCTGGCCAGCCCTCCCTCATCTCATCTGAGCC	2870		
Db	2826	TCACACCCCAAGAGCTGGCCAGCCCTCCCTCATCTCATCTGAGCC	2869		
RESULT 4	AY399041	1896 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	Mus musculus HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	AY399041				
VERSION	AY399041.1	GI:39755030			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.C., Adams,M.D. and Cargill,M.				
TITLE	Interfering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1896)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.C., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..1896				
gene	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	<1..>1896				
	/locus_tag="HCM0091"				
ORIGIN					
Query Match	61.2%, Score 1776, DB 9, Length 1896,				
Best Local Similarity	93.7%, Pred. No. 0,				
Matches 1776, Conservative	0; Mismatches 120; Indels 0; Gaps 0;				
123	ATGAGATCGGTGGCTTACTCCAGGCGCCGAGCCAGAGCTCCCTGGAGCTTCGACCTGGACC	182			

Dd	1	ATGAGTCGGTGGCTTTA	CTTCACAGCGCCGAGCCAGGCTCTCCCTGGGCTTCGCGCC	60
Qy	183	TCGGAGAGCCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAG		242
Dd	61	TCGGAGAGCCCGCGCGCTGGCGGAGGGGCTCATCAAGTCGCTAAACCTCTGATGAAG		120
Qy	243	AAGCAGCGCGTGAAGCGGCAACCATCAACAACAACCTTGCGGCAACCGTTCGAGTTCCTG		302
Dd	121	AAGCAGCGCGTGAAGCGGCAACCATCAACAACAACCTTGCGGCAACCGTTCGAGTTCCTG		180
Qy	303	GAGACGCTGGGCAAGGGCAACCTACGGGAAGGTGAAGAAGGCAAGAGAGCTCGGGGCGT		362
Dd	181	GAGACGCTGGGCAAGGGCAACCTACGGGAAGGTGAAGAAGGCAAGAGAGCTCGGGGCGT		240
Qy	363	CTGGTGGCCATCAAGTCATCAGAGAAAGCAAAATCAAGATGAGAGGATCTGTGCAC		422
Dd	241	CTGGTGGCCATCAAGTCATCAGAGAAAGCAAAATCAAGATGAGAGGATCTGTGCAC		300
Qy	423	ATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCAATCAT		482
Dd	301	ATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCAATCAT		360
Qy	483	GAACTGTTAGAAATAGCAGCAAGATTGGAATTGTCAATGAGAGTATCCAGCGAGCGAT		542
Dd	361	GAACTGTTAGAAATAGCAGCAAGATTGGAATTGTCAATGAGAGTATCCAGCGAGCGAT		420
Qy	543	CTGTATGATTTACATCAATGAGCGGCAACGGCTAGTGAAGCGGAGCCCAAGCAATTTCTTC		602
Dd	421	CTGTATGATTTACATCAATGAGCGGCAACGGCTAGTGAAGCGGAGCCCAAGCAATTTCTTC		480
Qy	603	CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGACGGAGTGTTCACCGAGATCTTC		662
Dd	481	CGACAGATCGTGTCTGCGCTGCACTACTGCGACAGACGGAGTGTTCACCGAGATCTTC		540
Qy	663	AAGTGGAAAAACATCCTCTTCTGATATGCCAATGAAAACATCAAGATTGCTGACTTTGGCCTTC		722
Dd	541	AAGTGGAAAAACATCCTCTTCTGATATGCCAATGAAAACATCAAGATTGCTGACTTTGGCCTTC		600
Qy	723	TCCAACCTGTACCAACAAGGCAAGTTCCTCACAACGTTCTGTGGAGCCCTCTCTACGCC		782
Dd	601	NN		660
Qy	783	TCGCTTGAATAGTCAACGGGAAACCCCTATGTGGGCCAGAGGTGACACAGTGTCTCTG		842
Dd	661	NN		720
Qy	843	GGCCTTCTCTGTATCATCTCTGTGTGATGGACCATATGCCCTTTTACACGGGACGATTCATAAA		902
Dd	721	GGCCTTCTCTGTATCATCTCTGTGTGATGGACCATATGCCCTTTTACACGGGACGATTCATAAA		780
Qy	903	ACACTGTGTGAAGCAATCATGTACGGGGCTTACCGTAGAGCGGCGCAAGCGTCCGATATGCC		962
Dd	781	ACACTGTGTGAAGCAATCATGTACGGGGCTTACCGTAGAGCGGCGCAAGCGTCCGATATGCC		840
Qy	963	TGTGGCCTGATCCGGTGGCTGTTAATGTGTGAACCCACCCGTCGGGCGCACATGAGAGAT		1022
Dd	841	TGTGGCCTGATCCGGTGGCTGTTAATGTGTGAACCCACCCGTCGGGCGCACATGAGAGAT		900
Qy	1023	GTAGCCAGTCAATTGTGTGGTCAACTGGGGTTTACCAACCGAGTCCGGGGAACAGAAAGCC		1082
Dd	901	GTAGCCAGTCAATTGTGTGGTCAACTGGGGTTTACCAACCGAGTCCGGGGAACAGAAAGCC		960
Qy	1083	CTGGTAGAGGATGGGACCCCTAATGGGTGACTTTGGGCGGGGCTTCATATGGGGACCTGTTTA		1142
Dd	961	CTGGTAGAGGATGGGACCCCTAATGGGTGACTTTGGGCGGGGCTTCATATGGGGACCTGTTTA		1020
Qy	1143	CGTGCCTCTTCGGGCGCCCTCTGTGAAGATGAGCCAGAGGTGTGCAGCTTCTTCAAGCAG		1202
Dd	1021	CGTGCCTCTTCGGGCGCCCTCTGTGAAGATGAGCCAGAGGTGTGCAGCTTCTTCAAGCAG		1080
Qy	1203	CACGTGCGGAGGTGGAACACTGTACTTGGGCTGAGCGGCAACATTCCTTAAAGAG		1262

Db 1081 CAGCTGCCGGAGGTGAGACATCTGCTGGGCTGAGCGGCAACATTTCTTAAGAG 1140

Qy 1263 TCCGGAAGAGATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGATACCTCT 1322

Db 1141 TCCGGAAGAGATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGATACCTCT 1200

Qy 1323 TCTGCGCTGGCAAGAGACCTTAAGCTTCCGAAAGGCAATTCAGAAAAAGTCTCT 1382

Db 1201 TCTGCGCTGGCAAGAGACCTTAAGCTTCCGAAAGGCAATTCAGAAAAAGTCTCT 1260

Qy 1383 ACCCTGTCAGGGGAGGTACAGAGACCTTACAGAACTCAGACCGGCTGATACCTCA 1442

Db 1261 ACCCTGTCAGGGGAGGTACAGAGACCTTACAGAACTCAGACCGGCTGATACCTCA 1320

Qy 1443 GGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1502

Db 1321 GGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380

Qy 1503 CAGCGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1562

Db 1381 CAGCGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440

Qy 1563 GCCAGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1622

Db 1441 GCCAGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

Qy 1623 CTCTCTCTCCAGCGGAGGAGCTTCTCAAACTCAATGGCAAGTCTCCCGACAGCCTTA 1682

Db 1501 CTCTCTCTCCAGCGGAGGAGCTTCTCAAACTCAATGGCAAGTCTCCCGACAGCCTTA 1560

Qy 1683 GAAAGCACTACCTCTGAGCACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1742

Db 1561 GAAAGCACTACCTCTGAGCACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620

Qy 1743 GCCCGCCGACCGCCCTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1802

Db 1621 GCCCGCCGACCGCCCTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680

Qy 1803 TTGACCAATTTGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1862

Db 1681 TTGACCAATTTGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

Qy 1863 GTGAGCACTGAGGAGGCTTGAAGCACTTCTGAGAGGCTGAGAGGCTGAGAGGCTG 1922

Db 1741 GTGAGCACTGAGGAGGCTTGAAGCACTTCTGAGAGGCTGAGAGGCTGAGAGGCTG 1800

Qy 1923 GAATCTTGGGGAGTACGCTGCTTTCTGTAAGAAGCTGCAAGAGGCTGCAAGCTTAC 1982

Db 1801 GAATCTTGGGGAGTACGCTGCTTTCTGTAAGAAGCTGCAAGAGGCTGCAAGCTTAC 1860

Qy 1983 AGACAAGCCCTGAGATCTGCTCAAGGCTGAGCTGA 2018

Db 1861 AGACAAGCCCTGAGATCTGCTCAAGGCTGAGCTGA 1896

RESULT 5
HSM801859
LOCUS HSM801859 3443 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp434U037 (from clone DKFZp434U037).
ACCESSION AL136891
VERSION AL136891.1 GI:12053280
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3443)
Kohrer, K., Beyer, A., Mewes, H. W., Weill, B., Amdt, C., Oesinger, A.,
Fodor, G., Han, M. and Wiemann, S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BMFZ (Biomedical Research Center) at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp434U037) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434U037
Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES
Source
Location/Qualifiers
1..3443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434U037"
/db_xref="taxon:9606"
/clone="DKFZp434U037"
/issue_type="testis"
/clone_id="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein"
1..3443
/gene="DKFZp434U037"
128..2014
/gene="DKFZp434U037"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB6825.1"
/db_xref="GI:12053281"
/db_xref="GOA:O9H093"
/translation="MESLIVPARBSPGPPSAARILAPLAGLILSPKIMKQAVRHH
KHNLRIRIEPLETLGKTKGKAKAESGRVLAITSIRDKIKEDDLNHRIRIE
IMSLNPHILALHEVENSSEKIVIVNEVSRGLDYIISRQQLSEREAFPRQIV
SAHYCHQNVVHEDLKENILIDANENIKIDAGLENLVHQGFLOFCEASPLVAP
EIVNGKEPTGPEVDSMSGLVLLVILVNGVFPDHDHKLIVKQISNGAYRPPPSDA
CGILRMILMNPTRATLEDAVLSHWNVGATVAGVDEAPHEGSGSDASLMSMD
WRRSSRPILENAKVCSFPKQARPGGSGTTPGLERHSLKSKENDMAQSLSDTA
DDIARHPKSNLKPXGILUKKVSASABGOEDPEPISPTIAPSGQAAPIIPKIGITL
KPRRESGYVSSPSSGSLDAGDVFSGDPEQKRPQASGLILHRKGLILKNGKF
SOTALELAFTTFSIDELAPRPLAPASPSGAVSDSILSSSFDOLDPERLPEP
PLRGCVSDNLTGLEBPPSGPSCLRRMRDPLGDSFSLTDCQEVATATYQALRVC
SKLT"

ORIGIN
Query Match 47.4%; Score 1374.2; DB 3; Length 3443;
Best Local Similarity 81.6%; Pred. No. 7.5e-313;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy 82 CTATGCTGCCGACCCCACTCACTGCGGCTCCCGGACCATGAGTGGCTTAC 141
Db 87 CTATGATTCCTCCGCGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 146
Qy 142 TCGAGCGCCGAGCGAGGCTCCCTGCGGCTCGCGCTGCGGCTGCGAGAGCGCCGCGCG 201
Db 147 CGCGGCGCTCGGCGCCCACTCTCGGCGGAGAGCTA-----GCCGCGCGC 194
Qy 202 TGCGGAGCGGCTCATGAGTGGCTTAACTCTGATGTAAGAGAGCGGCTGAGCGCGC 261
Db 195 TGCGGAGAGGCTATCAAGTCGCGCAAGCCCTTAATGAAGACAGCGGCTGAGAGCGCG 254
Qy 262 ACATCAACAACAACCTGCGGACCGGCTGAGAGTCTTGAAGACGCTGGGCAAGGCGCA 321
Db 255 ACACACCAACAACAACCTGCGGACCGGCTGAGAGTCTTGAAGACGCTGGGCAAGGCGCA 314
Qy 322 CCTACGGGAGAGTGAAGAGGACGAGAGAGCTGCGGCGCTGCTGAGGCTCATCAAGTCA 381
Db 315 CCTACGGGAGAGTGAAGAGGCGGAGAGAGCTGCGGCGCTGCTGAGGCTCATCAAGTCA 374
Qy 382 TCAGAAAAGACAAATCAAGATGAGAGATCTGCTGACATACGAGGAGGAGATTGAGA 441

FEATURES
Source Location/Qualifiers
1..1887
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1887
/locus_tag="HCM0091"

gene

ORIGIN

Query Match 43.6%; Score 1264; DB 9; Length 1887;
Best Local Similarity 78.5%; Pred. No. 7.4e-287;
Matches 1461; Conservative 0; Mismatches 380; Indels 21; Gaps 2;

QY 169 CTTCCGCTGCTGCTGAGAGCCGCGCTGCGGACGCGCTCATCAATGCTCTA 228
Db 35 CCACTCTCCGCGGCGAGAGCTAGCCGCGCTGCGGAGGCGTATCAAGTCGCCA 94

QY 229 AACTCTGATGAGAGAGAGCGGCTGAGAGCGGACCATCAAAACAACTGCGGACC 288
Db 95 AGCCCTTAATGAGAGAGAGCGGCTGAGAGCGGACCAACAACTGCGGACC 154

QY 289 GCTACGAGTTCCTGAGAGCGCTGAGAGCGGACCTACGCGAGGTGAGAGGCAAG 348
Db 155 GCTACGAGTTCCTGAGAGCGCTGAGAGCGGACCTACGCGAGGTGAGAGGCGG 214

QY 349 AGAGCTCGGCGCTGCTGCTGAGAGCGGACCATCAAGTCAAGAGAGAGAGATGAGC 408
Db 215 AGAGCTCGGCGCTGCTGCTGAGAGCGGACCATCAAGTCAAGAGAGAGATGAGC 274

QY 409 AGAGTCTGCTGACATACGAGAGAGATGAGATCATGCTTCACTCAACCCCA 468
Db 275 AAGATCTGATGACATACGAGAGAGATGAGATCATGCTTCACTCAACCCCA 334

QY 469 TCAATGCTCAATCAATGAGTGTGAGATGAGAGAGATGAGATGAGATGAGATGAG 528
Db 335 TCAATGCTCAATCAATGAGTGTGAGAGAGATGAGATGAGATGAGATGAGATGAG 394

QY 529 CCAAGCCGAGGAGATCTGATGATTAATCAATGAGAGCGGACGCTGAGAGCGGAGC 588
Db 395 CCAAGCCGAGGAGATCTGATGATTAATCAATGAGAGCGGACGCTGAGAGCGGAGC 454

QY 589 CCAAGCAATTTCTTCCGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
Db 455 CTAAGCAATTTCTTCCGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 514

QY 649 TTCAACCAAGATCTCAAGCTGAGAGAGATCTTCTGATGAGAGAGAGAGAGATG 708
Db 515 TTCAACCAAGATCTCAAGCTGAGAGAGATCTTCTGATGAGAGAGAGAGAGATG 574

QY 709 CTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTCTCTCAAGCTTCTGTGGA 768
Db 575 NNN 634

QY 769 GCCCTCTCTACGCTCGCTGAGATGATCAACGAGAGCCCTATGTGGGCCAGAGGTG 828
Db 635 NNN 694

QY 829 ACAGCTGTCTCTGAGGCTTCTCTCTGATCACTCTGCTGATGAGAGAGAGAGAG 888
Db 695 ACAGCTGTCTCTGAGGCTTCTCTCTGATCACTCTGCTGATGAGAGAGAGAGAG 754

QY 889 GAGAGATCAATAAAG 948
Db 755 GAGAGATCAATAAAG 814

QY 949 AGCGCTCGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
Db 815 AACCTCTGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 874

QY 1009 CCAACCTGAG 1068
Db 875 CCAACCTGAG 934

QY 1069 GGAACACGAG 1128
Db 935 GAG 994

QY 1129 TGGCGAGCTGCTGAG 1188
Db 995 TGGCGAGCTGCTGAG 1054

QY 1189 GCTTCTTCAAG 1248
Db 1055 GCTTCTTCAAG 1114

QY 1249 ATTCTTAAAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
Db 1115 ATTCTTAAAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1174

QY 1309 CTGAGATCACTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1368
Db 1175 CTGAGATCACTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234

QY 1369 AGAAAGATCTTCAAG 1428
Db 1235 AGAAAGATCTTCAAG 1294

QY 1429 TGCTGATCACTTCAAG 1488
Db 1295 TGCTGATCACTTCAAG 1345

QY 1489 TTAAGAGCTTCAAG 1548
Db 1346 TTAAGAGCTTCAAG 1405

QY 1549 GGAAGATCTTCAAG 1608
Db 1406 GGAAGATCTTCAAG 1465

QY 1609 CACAGGCTTCAAG 1668
Db 1466 CACAGGCTTCAAG 1525

QY 1669 CCGGACAGCTTCAAG 1728
Db 1526 CCGGACAGCTTCAAG 1585

QY 1729 CTTCCAGATCTTCAAG 1788
Db 1586 CTTCCAGATCTTCAAG 1645

QY 1789 TGTCTCTGAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1848
Db 1646 TGTCTCTGAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1705

QY 1849 GGGGCTGTGTCTGAG 1902
Db 1706 GGGGCTGTGTCTGAG 1765

QY 1903 -----GTGAG 1956
Db 1766 GAGGCTCTGAG 1825

QY 1957 ACTGCAAG 2016
Db 1826 ACTGCAAG 1885

QY 2017 GA 2018
Db 1886 GA 1887

RESULT 7
AY39040 1887 bp DNA linear GSS 12-DEC-2003
LOCUS AY39040
DEFINITION Pan troglodytes HCM0091 gene, VIRUTAL TRANSCRIPT, partial sequence,

genomic survey sequence.
AY399040 GI:3975029
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
REFERENCE
AUTHORS
1 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
AUTHORS
2 (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submisison
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..1887
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1887
/locus_tag="HCM0091"
ORIGIN
Query Match 32.8%; Score 951.4; DB 9; Length 1887;
Best Local Similarity 60.1%; Pred. No. 4.2e-213;
Matches 1119; Conservative 0; Mismatches 722; Indels 21; Gaps 2;
QY 169 CCTCCGCGCTGAGCGCTGCGAGAGCGCGCGCTGCGGAGCGGGCTCATCAAGTCGCGTA 228
DB 35 CCACTCCCTCGCGCGCGAGAGCTAGCCGCGCGCTGCGGAGCGGCTGATCAAGTCGCCA 94
QY 229 AACCTCTGATGAAGACGAGCGGTGAAGCGGCACATCAACAACAACTGCGGAC 288
DB 95 AGCCCCATATGAAGAACAGCGCGGTGAAGCGGCACACCAACAACAACCTGCGGAC 154
QY 289 GCTACGAGTCTCTGAGAGCGCTGCGGAGCGGCACCTACGGAAGGTGAAGAGCGACGAG 348
DB 155 GCTACGAGTCTCTGAGAGCGCTGCGGAGCGGCACCTACGGAAGGTGAAGAGCGCGG 214
QY 349 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCAATCCGGAAGACAAATCAAAATGAGC 408
DB 215 AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCAATCCGGAAGACAAATCAAAATGAGC 274
QY 409 AGAGTCTGTGCAATACGAGGAGATTGAGATCATGTCTTCACTCAACCCGCCACA 468
DB 275 AAGATCTGATGCAATACGAGGAGATTGAGATCATGTCTTCACTCAACCCGCCACA 334
QY 469 TCAATGTCATTCATGAAGTGTGTAAGATAGACGAAGATTGTATGTCATGAGATAG 528
DB 335 TCAATGTCATTCATGAAGTGTGTAAGATAGACGAAGATTGTATGTCATGAGATAG 394
QY 529 CCAAGCGAGCGATCTGTATGATTAATCATAGTACGCGCAACGCTGAGTGAAGCGGACG 588
DB 395 NNN 454
QY 589 CCAAGCATTTCTTCCGACAGATCGTGTCTGCCCTGCACTAGCCACAGAACGGATCG 648
DB 455 NNN 514
QY 649 TTCAACGAGATCTCAAGCTGAAACATCTCTTAGATGCCATGAAACATCAAGATTG 708

DB 515 TCACCGAGATCTCAAGCTGAGAACATCTCTTGATGCGCAATGGAAATATCAAGNNN 574
QY 709 CTGACTTTGGCGCTCTCCAACTGTACCAAGGCAAGTTCTTCAGACGTTCTGTGAGA 768
DB 575 NNN 634
QY 769 GCCCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCTATGTGGGCCAGAGTGG 828
DB 635 NNN 694
QY 829 ACAGCTGTCTCTGGGGGTTCTCTGTACATCTGTGTGATGGACATGCGCTTTGACG 888
DB 695 NNN 754
QY 889 GGCAGATCATAAACACTGATGAGCAAAATCAGTAAACGGGGCTTACCGTAGCGGCCA 948
DB 755 NNN 814
QY 949 AGCCGTCGATGCTGTGGCTGTATCCGTTGCTGTTAATGTGAACCCACCGCTGGG 1008
DB 815 NNN 874
QY 1009 CCAACTGAGAGATGTAGCAATGTAGTATGTTGGTCACTGGGGTTACACACCGAGTGG 1068
DB 875 CCACTGAGAGATGTAGCAATGTAGTATGTTGGTCACTGGGGTTACACACCGAGTGG 934
QY 1069 GGGAAACAGAAAGCTGCTGAGAGGTGGGACCTAGTGTGACTTTGGCGGGCTTCCA 1128
DB 935 GAGAGCAGAGAGGCTCCGCAATGAGGTGGGACCTGCAATGACTGTCTCGCGCTCCG 994
QY 1129 TGGGGAATGTTACGTGCTCTCTGCGGCCCTCTCTGAGAAATGAGCAAGTGTGCA 1188
DB 995 TGGTGAATGTTACGTGCTCTCTCTGCGGCCCTCTCTGAGAAATGAGCAAGTGTGCA 1054
QY 1189 GCTTCTTCAAGCAGACAGTCCCGGAGGTGGAAGCACTGATCTGGGCTGAGGCGGAC 1248
DB 1055 GCTTCTTCAAGCAGACAGTCCCGGAGGTGGAAGCACTGATCTGGGCTGAGGCGGAC 1114
QY 1249 ATTCTCTTAAGAGTCCCGAAAGAGATACATAGTCTCAAAATCTGCAAGTGAACCCG 1308
DB 1115 ATTCTCTTAAGAGTCCCGAAAGAGATACATAGTCTCAAAATCTGCAAGTGAACCCG 1174
QY 1309 CTGAGATACCTCTCTCTGCGCTTGGCAAGACGCTTAAAGCTTCCGAAAGGCAATTCTA 1368
DB 1175 CTGATGACATCTGCGCATCTGCGTGGCAAGACGCTTAAAGCTTCCGAAAGGCAATTCTA 1234
QY 1369 AGAAAGTCTCTTACCTCTCTGAGGAGGTATACAGAGGACCTTCAGGAACTCAGACCG 1428
DB 1235 AGAAAGTCTCTTACCTCTCTGAGGAGGTATACAGAGGACCTTCAGGAACTCAGACCG 1294
QY 1429 TGCCTGATCTCAAGGAGCACTGTCTCTGTATCTCTGCTCCCAAGGAAAGGCAATCC 1488
DB 1295 TCCCTGAGACCCAGGAGGAGGCTGCGC-----CCCTGCTCCCAAGGAGGCAATCC 1345
QY 1489 TTAAGAGTCTGACAGCGGTAATCTGTTACTTCTCTTCCAGAGCCGAGGAGTCTG 1548
DB 1346 TCAAGAGCCCAACAGCGGAGGTCTGACTACTCTCTCCGAGCCGAGGAGTCTG 1405
QY 1549 GGGAACTCTTAAGACGCAAGTGTGTGTTGTAGTGGGACCCCGTGGAGCAAGTCTC 1608
DB 1406 GGGAGCTCANNAGACGAGCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1465
QY 1609 CACAGGCTTCAAGGCTCTCTCTCAACGCAAGGCAATTCACAATCAATGGAAGTCT 1668
DB 1466 NNN 1525
QY 1669 CCGCAGAGCTTGAAGGACATACCTAGACCTTTGGCTCCCTGACCAATGCGCT 1728
DB 1526 NNNNGACAGCTTGGAGCTCGGGCCCCACACACTTGGCTCTTGGAGTGAAGTGGCCC 1585
QY 1729 CTTCCTCTCTGAGCGCCGCGCCCTCAAGGGCTGTGAGTGAAGACAGCTCC 1788

Db 1586 CATCTGCCCCCTGGCCCGGCGAGCCGACCTCAAGGGCTGTGAGGAGGACAGCATCC 1645
Qy 1789 TGCTCTCCGAGTCTTGTGACCAATTTGACCTTGGCCGAAAGTCTCCGAAACCCCACTGA 1848
Db 1646 TGCTCTGAGTCTTGTGACCAAGCTTGTGCTTGGCCGAAAGTCTCCGAAACCCCACTGC 1705
Qy 1849 GGGGCTGTGTCTGTGTGACCAAGCTTGAAGGGGCTTGAACAGCCTCCCTCAGAG----- 1902
Db 1706 GGGGCTGTGTCTGTGTGACCAAGCTTGAAGGGGCTTGAAGGGGCTTGAAGGGGCTTGA 1765
Qy 1903 -----GTCTGAAGCAGTGTGTGAGGAGTCTTGGGGGATAGCTGCTTTTCTGTGACG 1956
Db 1766 GAACTCTGCTGAGGCGCTGTGGCGGAGGATCTTGTGGGGGAGCAGCTGCTTTTCCCTGACG 1825
Qy 1957 ACTGCCAAGAGGTGACTGACCTGACCTTGAAGAGAGGCTTGAAGTCTGCTCAAGCTCAGCT 2016
Db 1826 ACTGCCAAGAGGTGACAGCGACTTACCGGACGACCTGAGGGGTCTGTCAAGCTCACT 1885
Qy 2017 GA 2018
Db 1886 GA 1887

RESULT 8
CR749209 7004 bp mRNA linear HTC 19-AUG-2004
LOCUS Homo sapiens mRNA; cDNA DKFP686F01113 (from clone DKFP686F01113).
DEFINITION CR749209
ACCESSION CR749209.1 GI:51476151
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 7004)
Ostenwald, B., Obermaier, B., Deutscher, S., Schallp, A.,
Mewes, H.W., Weil, B., Amlid, C., Osanger, A., Fobo, G., Han, M., and
Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by Medigenomix (Martinsried/Germany) within the cDNA
Sequencing consortium of the German Genome Project.
This clone (DKFP686F01113) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFP686F01113
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
Location/Qualifiers
1..7004
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686F01113"
/issue_type="fetal kidney"
/clone_lib="686 (synonym: hlc3). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stages="fetal"
/note="hypothetical protein, N-terminus truncated, not
fully spliced"
1..7004
/gene="DKFP686F01113"
/gene="DKFP686F01113"
/gene="DKFP686F01113"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18066.1"
/db_xref="GI:51476152"
/translation="IADPGLSNLYHGKFLQTFQSGPLVYASPEIVNGKPYTPYDVS"

SLGLVLYILVHGTMPEDGHDHKLIVQISNGVAREPKPSDAGLIMLVNPTERRA
TLBDAVSHMWNMGYATRVGEQAEHFGHGRGSDSASADWLRSRPLLENKARY
CSFPEKJAPGSGSTTPELRQSLKSKRKNMDAQSHSDIADTADTPRPGKSNLKP
GILKKKYSABGVQEDPELSPAPAGVAPLLPKGGLIKKROESGYSPEPS
ESSELIDAGAGVPVSDPEKRPQASGLILHRGKILKLNKESQTLIELAPTFGSL
DELAPPEPLARASRPSGAVESDILSSESDQDLPLRLPPEPLRGCVSDNLTGLES
PSEBGPSCILRWRMDPLGDSFSLDQCEVATYVRLALRVCSKLTP"

ORIGIN

Query Match 32.2%; Score 934.4; DB 3; Length 7004;
Best local similarity 80.6%; Pred. No. 5.7e-209;
Matches 1137; Conservative 0; Mismatches 251; Indels 22; Gaps 3;

Qy 695 AATCATCAAGATTCTGACTTTTGCCCTCTCCAACTGTACCAAAAGCAAGTCTCTCA 754
Db 4153 AATCTCGAATGCTGACTTCCGCTCTCCAACTGTACCAAAAGCAAGTCTCTCA 4212
Qy 755 GACGTTCTGTGGAGCCCTCTTACGCTCGCTGAGATGTCACAGGAAAGCCCTATGT 814
Db 4213 GACATTCGTGGAGGCCCTCTATGCTCGCCAGAGATTGTCAATGGAGCCCTACAC 4272
Qy 815 GGGCCCAAGGTGACAGCTGTGTCTGTGGGCTTCTCTGTACATCTGTGTGACATGAC 874
Db 4273 AGGCCAAGGTGACAGCTGTGTCTGTGGGCTTCTCTGTACATCTGTGTGACATGAC 4332
Qy 875 CATGCCCTTTGACGGGAGGATCATAAACACTGATGAAGCAATCAAGTACGGGCTTGA 934
Db 4333 CATGCCCTTTGATGGGATGACATTAAGTCTTGTGAACAGATCAGCAACGGGCTTGA 4392
Qy 935 CCGTGAACCGGCCAAGCCGCTCGATGCTGTGGCTGATCCGATGCTTTAATGTGAA 994
Db 4393 CCGGAGACCACTTAACCCCTGTATGCTGTGCTGATCCGATGCTTTAATGTGAA 4452
Qy 995 CCCCACCCGTCGGGCCCACTGAGAGATGACCAATGATTTGGTGTCACTGGGCTTA 1054
Db 4453 CCCCACCCGTCGGGCCCACTGAGAGATGACCAATGATTTGGTGTCACTGGGCTTA 4512
Qy 1055 CACCAAGGATGGGGGAAACAGAGAGCCGCTGAGAGGAGGAGCCTTATGATGACTT 1114
Db 4513 CCGCACCCGATGGAGAGAGAGAGGCTCCGATAGAGGATGAGCCTTATGATGACTT 4572
Qy 1115 TGCCGGGCGCTCATAGCGGAGCTGTGATAGTGTGCTCGGCCCTCTCTGAGAAATG 1174
Db 4573 TGCCGGGCGCTCATAGGCTGTGATAGTGTGCTCGGCCCTCTCTGAGAAATG 4632
Qy 1175 AGCCAAAGTGTGACGCTTTCTCAAGCAGCAGCTGCGGAGGTGAGAACATGACTGG 1234
Db 4633 GGCACAAAGTGTGACGCTTTCTCAAGCAGCAGCTGCGGAGGTGAGAACATGACTGG 4692
Qy 1235 GCTGGAGCGGCAATCTCTTAAGAGTCCCGAAAGAGATGATGCTCAAAATCT 1294
Db 4693 CTTGAGCGCGCAGATTTGCTCAAGAGTCCCGAAAGAGATGATGCTCAAAATCT 4752
Qy 1295 GCAAGTGAACCCGCTGAGATATCTCTTCTCGCCCTGCAAGAGCAGCTTAAAGTCTCC 1354
Db 4753 CCAAGTGAACCCGCTGATATGATATGCTGCTGCTGCAAGAGCAGCTTAAAGTCTCC 4812
Qy 1355 GAAAGGCAATTTCAAGAAAAAGTCTTACTCTGTCAGGGAGGTACAGAGAACCTTCA 1414
Db 4813 AAGAGGCAATTTCAAGAAAAAGTGTGACGCTTGTGAGAAAGGGTACAGAGAACCTTCC 4872
Qy 1415 GGAATCAGACCGGTGCTGATATCTCAAGGAGCAGCTGTCCCTGCTGATATCTCTGCTCC 1474
Db 4873 GGAATCAGACCGGTGCTGATATCTCAAGGAGCAGCTGTCCCTGCTGATATCTCTGCTCC 4923
Qy 1475 AAGGAAAGGCAATCTTAAGAGTCTGACAGCGTGAATCTGTACTACTCTCTCCAGA 1534
Db 4924 CAAGAAGGCAATTTCAAGAAAGCCCGACAGCGGAGGTCTGTGCTACTACTCTCTCCAGA 4983
Qy 1535 GCCCAGGAGTCTGTGGGAACTTTAGAGCGCAGAGATGTTTGTGAGTGGGAGCCCGGT 1594
Db 4984 GCCCAGGAGTCTGTGGGAGCTTTGAGAGCGAGGAGCAGTGTGAGTGGGAGTCCCA 5043

QY 1595 GAGCAGAAAGTCTCCACAGGCTTCAGGGCTCCTCCTCCACGCAAGGCAATTCCTCAACT 1654
DB 5044 GAGCAGAAAGCTCCGCAAGCTTCAGGGCTCCTCCTCCACGCAAGGCAATTCCTCAACT 5103
QY 1655 CAATGCAAGTCTCCGCAAGCTTCAGGGCTCCTCCTCCACGCAATTCCTCCTCCCT 1714
DB 5104 CAATGCAAGTCTCCGCAAGCTTCAGGGCTCCTCCTCCACGCAATTCCTCCTCCCT 5163
QY 1715 GAGCAGAAAGTCTCCACAGGCTTCAGGGCTCCTCCTCCACGCAAGGCAATTCCTCAACT 1774
DB 5164 GAGCAGAAAGCTCCGCAAGCTTCAGGGCTCCTCCTCCACGCAAGGCAATTCCTCAACT 5223
QY 1775 TGAGCAGAGATCTCTGCTCCGAGTCTTCGCAATTCGCAATTCGCAATTCGCAATTCG 1834
DB 5224 CGAGCAGAGATCTCTGCTCCGAGTCTTCGCAATTCGCAATTCGCAATTCGCAATTCG 5283
QY 1835 CGAAACCCCACTGAGGGGCTGTGTCTGTGCAACAACCTGAGGGGCTTGAAGAGCCCTCC 1894
DB 5284 AGAGCCCACTGAGGGGCTGTGTCTGTGCAACAACCTGAGGGGCTTGAAGAGCCCTCC 5343
QY 1895 CTGAGAG-----GTGAGAGCAGTGTGTGAGAGATTCCTTGGGGGATAGCTG 1942
DB 5344 CTGAGAGGGCTGTGAGAGCTGCTGAGAGGCTGTGAGAGATTCCTTGGGGGATAGCTG 5403
QY 1943 CTTTCTCTGACAGATGCGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2002
DB 5404 CTTTCTCTGACAGATGCGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5463
QY 2003 CTCAAGCTCAGTGAAGAGAGAGATGTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2062
DB 5464 CTCAAGCTCAGTGAAGAGAGAGATGTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 5522
QY 2063 TTGAGAGAGAACCTGTGCTGAGATTCCTCC 2092
DB 5523 GCTGCTTGCACCCGAGAGAGATGCTTC 5552

RESULT 9
LOCUS B1653092 746 bp mRNA linear EST 12-SEP-2001
DEFINITION 603300616F1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:5341299 5',
mRNA sequence.
ACCESSION B1653092
VERSION B1653092.1 GI:15567328
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 746)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennigshausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LMNL1866 row: p column: 04
High quality sequence stop: 746.
Location/Qualifiers
1. .746
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5341299"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"

/lab_host="DH10B"
/clone_lib="NCI CGAP Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennigshausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

Query Match 24.8%; Score 721; DB 4; Length 746;
Best Local Similarity 98.0%; Pred. No. 8e-159;
Matches 730; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1787 CTTCTCTCCGAGTCTTTTGAACAATGAACTGCTGAAAGCTTTCGCAAGCCCACT 1846
DB 2 CCAAGCGTCCGAGCCACGCTCCGTAATGAGCTTGCCTGAAGCTTCCGAAACCCACT 61
QY 1847 GAGGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCTCCTCAAGAGTCT 1906
DB 62 GAGGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCTCCTCAAGAGTCT 121
QY 1907 GAAAGCAGTGTGAGAGAAATCTTGGGGATAGCTGCTTCTGACAGACTGCAAGA 1966
DB 122 GAAAGCAGTGTGAGAGAAATCTTGGGGATAGCTGCTTCTGACAGACTGCAAGA 181
QY 1967 GGTGATCTGACGCTTACAGACAGACGCTTGAAGATCTGCTCAAGAGTGAAGAGGA 2026
DB 182 GGTGATCTGACGCTTACAGACAGACGCTTGAAGATCTGCTCAAGAGTGAAGAGGA 241
QY 2027 GATGAGCCCTAGATGAGGGGATAGGCTGAGAGGCTTGAAGAGAACTTGGGTCGAT 2086
DB 242 GATGAGCCCTAGATGAGGGGATAGGCTGAGAGGCTTGAAGAGAACTTGGGTCGAT 301
QY 2087 TCCCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2146
DB 302 TCCCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
QY 2147 ATGAGAGAAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2206
DB 362 ATGAGAGAAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
QY 2207 TGGGGTGGCTGAAAGGCTTACCTCTTGTGACATGATGATGATGATGATGATGATGATGAT 2266
DB 422 TGGGGTGGCTGAAAGGCTTACCTCTTGTGACATGATGATGATGATGATGATGATGATGAT 481
QY 2267 ACCCTGTTCTGAGCTGACCTTCACTAAGTTTCTGCTCAATCAACCAAGAGGTTAG 2326
DB 482 ACCCTGTTCTGAGCTGACCTTCACTAAGTTTCTGCTCAATCAACCAAGAGGTTAG 541
QY 2327 AACCTGACTTCTGAGAGGTAATGTGTAGTGAATGATGATGATGATGATGATGATGATGAT 2386
DB 542 AACCTGACTTCTGAGAGGTAATGTGTAGTGAATGATGATGATGATGATGATGATGATGAT 601
QY 2387 TCTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
DB 602 TCTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 2447 TGACTTCATCTCAAGGAGCAAGATGCTTGAAGCCCATCTTATGATCTAGAGACTTGA 2506
DB 662 TGACTTCATCTCAAGGAGCAAGATGCTTGAAGCCCATCTTATGATCTAGAGACTTGA 721
QY 2507 ACCTTGAAGCTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2531
DB 722 ACCTTGAAGCTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746

RESULT 10
LOCUS CA319312 795 bp mRNA linear EST 09-JUL-2003
DEFINITION UT-M-FMO-cbz-i-16-0-UT-r1 NIH BMAP_FMO Mus musculus cDNA clone
IMAGE:6816569 5', mRNA sequence.
ACCESSION CA319312

VERSION CA319312.1 GI:24537436
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 795)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: PYX-5.
 Location/Qualifiers
 1..795
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6816569"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_id="NIH BMAP FMO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 24.8%; Score 720; DB 6; Length 795;
 Best Local Similarity 96.8%; Pred. No. 1,4e-156;
 Matches 765; Conservative 0; Mismatches 0; Indels 25; Gaps 2;

QY 43 GCGGCTGTGACCTCTGAGCCGCGGCTCAGGCGGCTGCTACTGCTGCCGACCACTC 102
 DB 1 GCGGCTGTGACCTCTGAGCCGCGGCTCAGGCGGCTGCTACTGCTGCCGACCACTC 60
 QY 103 CACCTGCGGAGTCCCGGACCATGAGAGTGGTGGCTTACTCAGCGCCGAGGAGGCTC 162
 DB 61 CACCTGCGGAGTCCCGGACCATGAGAGTGGTGGCTTACTCAGCGCCGAGGAGGCTC 120
 QY 163 CCGTGGGCTCCGCGGCTCGGAGAGCGCGGCTGCGGAGCGGCTCATCAAGT 222
 DB 121 CCGTGGGCTCCGCGGCTCGGAGAGCGCGGCTGCGGAGCGGCTCATCAAGT 180
 QY 223 GCGCTAAACCTCTGATGAGAGAGCGGCTGAGCGGACCATCAACCAACCTGTC 282
 DB 181 GCGCTAAACCTCTGATGAGAGAGCGGCTGAGCGGACCATCAACCAACCTGTC 240
 QY 283 GGCACCGGCTACGATGCTCTGAGAGCGTGGGCAAGGGGACCTACGGGAAAGTGAAGAG 342
 DB 241 GGCACCGGCTACGATGCTCTGAGAGCGTGGGCAAGGGGACCTACGGGAAAGTGAAGAG 300

QY 343 CACGAGAGAGCTCGGGGCGTCTGTGGCCATCAAGTCATCAGGAAGAAATCAAG 402
 DB 301 CACGAGAGAGCTCGGGGCGTCTGTGGCCATCAAGTCATCAGGAAGAAATCAAG 360
 QY 403 ATGAGCAGAGTCTGCTGCAATACAGGAGAGATTTAGATCATCTCTCAACCAAC 462
 DB 361 ATGAGCAGAGTCTGCTGCAATACAGGAGAGATTTAGATCATCTCTCAACCAAC 420
 QY 463 CCCACATCATTTGCCATCATGA-----GTGTTGAGATA 498
 DB 421 CCCACATCATTTGCCATCATGAAGTGGGAGATCAGCTTAGTACTGTGTTAGATA 480
 QY 499 GCAGCAGATTGTGATTTGATGATGAGAGTATGCGAGCGAGATCTGTATGATTAATCA 558
 DB 481 GCAGCAGATTGTGATTTGATGATGAGAGTATGCGAGCGAGATCTGTATGATTAATCA 540
 QY 559 GTGAGCGGCGACCGGCTGATGAGAGCGGAGCCGAGGCAATTTCTTCCGACAGATCGTCTG 618
 DB 541 GTGAGCGGCGACCGGCTGATGAGAGCGGAGCCGAGGCAATTTCTTCCGACAGATCGTCTG 600
 QY 619 CCCTGCACTACTGCGACAGAGGGGATCGTTACCGAGATCTCAGGCTGGAAGAAATCC 678
 DB 601 CCCTGCACTACTGCGACAGAGGGGATCGTTACCGAGATCTCAGGCTGGAAGAAATCC 660
 QY 679 TTCTAGATGCGCAATGAGAAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACACA 738
 DB 661 TTCTAGATGCGCAATGAGAAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACACA 720
 QY 739 AAGCAGAGTCTCTCAAGCGTTCTGTGGAGCGCTCTCTACGCTCGCTGAGATATGTA 798
 DB 721 AAGCAGAGTCTCTCAAGCGTTCTGTGGAG-CCTCTCTACGCTCGCTGAGATATGTA 779
 QY 799 ACGGAGAGCC 808
 DB 780 ACGGAGAGCC 789

RESULT 11
 CF729223 727 bp mRNA linear EST 09-OCT-2003
 LOCUS U1-M-HD0-clc-b-24-0-UT.r1 NIH BMAP HD0 Mus musculus cDNA clone
 DEFINITION IMAGE:30615119 5', mRNA sequence.
 ACCESSION CF729223
 VERSION CF729223.1 GI:37603391
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 727)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: PYX-5.
 Location/Qualifiers
 1..727
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30615119"
 /tissue_type="whole eye"

/dev stage="embryo.12.5.13.5.14.5 dpc"
/lab host="DH10B (T1 phage resistant)"
/clone lib="NTH_BMAP_HD0"
/note="Organ: Eye; Vector: pYX-Abs; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Abs vector. The library tag
sequence located between the Not I site and the polyA tail
is TTAATGAAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
Developing Mouse Nervous System", supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 24.6%; Score 714; DB 7; Length 727;
Best Local Similarity 99.6%; Pred. No. 3.6e-157;
Matches 725; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1828 GTCTTCCGAAACCCCACTGAGGAGCTGTGTCTGTGAGACACCTGAGGAGGCTTGAAC 1887
1 GTCTTCCGAAACCCCACTGAGGAGCTGTGTCTGTGAGAC-ACCTGAGGAGGCTTGAAC 59
1888 AGCTCCCTCGAAGAGTCTGAAGAGATGTGTGCGAGAAATCCTTGGGGGATAGCTGCTTTT 1947
60 AGCTCCCTCGAAGAGTCTGAAGAGATGTGTGCGAGAAATCCTTGGGGGATAGCTGCTTTT 119
1948 CTCTGACGAGCTGCAAGAGGTGACTGAGGCTTACAGACAGCCCTGAGGATTCGCTCAA 2007
120 CTCTGACGAGCTGCAAGAGGTGACTGAGGCTTACAGACAGCCCTGAGGATTCGCTCAA 179
2008 AGCTCAGCTGAGAGGAGATGTGTGCTGTGTGAGGCTTGAAGAGGCTTGAAGAGGCTTGA 2067
180 AGCTCAGCTGAGAGGAGATGTGTGCTGTGTGAGGCTTGAAGAGGCTTGAAGAGGCTTGA 239
2068 GAGAAACCTTGGGTCGATTTCTCCAGTGAATGATGATCAATCAAGGCTTGAAGCTTGA 2127
240 GAGAAACCTTGGGTCGATTTCTCCAGTGAATGATGATCAATCAAGGCTTGAAGCTTGA 299
2128 GCGTACCTGAACTGAAAGATGAGAAATGCAATGATGAGAAAGGAAAGGAAACCTT 2187
300 GCGTACCTGAACTGAAAGATGAGAAATGCAATGATGAGAAAGGAAAGGAAACCTT 359
2188 TGCTGCCGAGTGTATAGTGGGTCGCTGAGAGTGCCTACTCTTTTGCCATGAGT 2247
360 TGCTGCCGAGTGTATAGTGGGTCGCTGAGAGTGCCTACTCTTTTGCCATGAGT 419
2248 GTCAACCATGACATTTCCACCCCTGTTCTGTGCTGACCTTCAATGATTTCTGTTTC 2307
420 GTCAACCATGACATTTCCACCCCTGTTCTGTGCTGACCTTCAATGATTTCTGTTTC 479
2308 CATCAACCAACGAGGTTTGAACCTGACTTCTGAGGAGTATATGTATGATGATGATGAT 2367
480 CATCAACCAACGAGGTTTGAACCTGACTTCTGAGGAGTATATGTATGATGATGATGAT 539
2368 ATTAGAGAGGAAACAGCTCTGTGTTTCATCTGTGCTGTGCTGATCTCAAGACCTGG 2427
540 ATTAGAGAGGAAACAGCTCTGTGTTTCATCTGTGCTGTGCTGATCTCAAGACCTGG 599
2428 GAAGACTCGGACCCCTGTTGACTTCACTCAAGGGGACAGATGCCCCCTGAGACCCATC 2487
600 GAAGACTCGGACCCCTGTTGACTTCACTCAAGGGGACAGATGCCCCCTGAGACCCATC 659
2488 TTGATCTCAGAGCTTGAACCTTGAAGCTGTTCTTGTAGTACCCGAGATGTGATGATGCT 2547
660 TTGATCTCAGAGCTTGAACCTTGAAGCTGTTCTTGTAGTACCCGAGATGTGATGATGCT 719
2548 CTGTTTCT 2555

Db 720 CTGTTTCT 727

RESULT 12
BG915967
LOCUS
DEFINITION
602815231F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4937647 5',
mRNA sequence.
ACCESSION
BG915967
VERSION
BG915967.1 GI:14296443
KEYWORDS
SOURCE
MUS musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strauberg, Ph.D.
Email: ggaubs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10871 row: m column: 08
High quality sequence stop: 695.
Location/Qualifiers
1. 992
/organism="Mus musculus"
/mol type="mRNA"
/strain="NMRI"
/db xref="taxon:10090"
/clone="IMAGE:4937647"
/tissue="tumor, gross tissue"
/dev stage="5 months"
/lab host="DH10B"
/clone lib="NCI CGAP Mam4"
/note="Organ: mammary; pCMV-SPORT6; Site_1: SalI;
Site_2: NotI. Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

FEATURES

source

ORIGIN

Query Match 24.1%; Score 700.2; DB 4; Length 992;
Best Local Similarity 92.2%; Pred. No. 6.9e-154;
Matches 857; Conservative 0; Mismatches 60; Indels 12; Gaps 11;

1931 GGGGAGTAGTGTCTTCTCTGACAGCTGCCAAGAGTATCTGCGAGCTTACAGACAGC 1990
13 GGGGAGTAGTGTCTTCTCTGACAGCTGCCAAGAGTATCTGCGAGCTTACAGACAGC 72
1991 CCTAGAAATCTGCTCAAGAGCTCAGTGAAGAGGAGATGTGTCCTTGAATGAGTATGAG 2050
73 CTTAGAAATCTGCTCAAGAGCTCAGTGAAGAGGAGATGTGTCCTTGAATGAGTATGAG 132
2051 CTCTGAGAGGTTTGCAGAGAAACCTTGGTGCAGATTCTTCAAGTAAATGAGTATCA 2110
133 CTCTGAGAGGTTTGCAGAGAAACCTTGGTGCAGATTCTTCAAGTAAATGAGTATCA 192
2111 AGGGCTTACGTCTGACAGCTGAGTGAACCTGAAAGATGAGAAATTCGATGATGTGG 2170
193 AGGGCTTACGTCTGACAGCTGAGTGAACCTGAAAGATGAGAAATTCGATGATGTGG 252
2171 AAAGAAATGGAAACCTTGTGACCGAGTGTATAGTGGGAGGCTGAGAGTGCCTAC 2230
253 AAAGAAATGGAAACCTTGTGACCGAGTGTATAGTGGGAGGCTGAGAGTGCCTAC 312

QY 2231 TCCTTGTGCGCATGAGTGTCAACCCATGACATTTCCACCCCTGTTCTCTGCGTGCACCTTC 2230
 DB 313 TCCTTGTGCGCATGAGTGTCAACCCATGACATTTCCACCCCTGTTCTCTGCGTGCACCTTC 372
 QY 2291 ACATAGTTTCTGTTTTCATTAACCAACGAGGTATGAAACCTGACTTCTCTGGAGGTA-A 2349
 DB 373 ACATAGTTTCTGTTTTCATTAACCAACGAGGTATGAAACCTGACTTCTCTGGAGGTA-A 431
 QY 2350 TGTGTAGTGCATGATTAATTATAGAGAGAAAGACCTGTTTTCATCTCTGCTGCTG 2409
 DB 432 TGTGTAGTGCATGATTAATTATAGAGAGAAAGACCTGTTTTCATCTCTGCTGCTG 491
 QY 2410 TGCATCTCAAGACCTGAGAGACTCGGACCGCTGTTGACTTCATCTCAAGGAGACAG 2469
 DB 492 TGCATCTCAAGACCTGAGAGACTCGGACCGCTGTTGACTTCATCTCAAGGAGACAG 551
 QY 2470 ATGCCCTCGAAGCCCATCTTATGATCTGAGAGACTTGAACCTTGAAGCTGTTCTAGTACC 2529
 DB 552 ATGCCCTCGAAGCCCATCTTATGATCTGAGAGACTTGAACCTTGAAGCTGTTCTAGTACC 610
 QY 2530 CAGATGTGGA-TGGATGCTGTTTCTCAGGCGCAAGGAGACCTGAAGTGTGCTGACTAT 2588
 DB 611 CAGATGTGGA-TGGATGCTGTTTCTCAGGCGCAAGGAGACCTGAAGTGTGCTGACTAT 670
 QY 2589 TTATTTTGTGATCTCACTTCTGTTTGTGTT--TTGTTGTTGTTGTTGTTTGTG 2646
 DB 671 NTATCTCCCGTATCTCACTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 730
 QY 2647 TTTTAAAGTAAATTTTGTCT-GCTTCAATTAATGATGATGCTGTTCTGAGAACTGCAC 2705
 DB 731 TCTTAAGTGACATCTTGTGCTGCTTCACTGATGAAGTGTGTTGCTGAGG-ACCTCAC 789
 QY 2706 TGTGCCATGAGTTTATGTAAGAGAGATTTGGCAATGATG-TCCCTTATTTCAAG 2764
 DB 790 TGTGCC-CTGAAGTGTAGTGTGCGGAAAGTTTGGCATGATGTTCCCTTATTTCAAG 848
 QY 2765 GGGGTGAGGGGCTTTTCA-AATGATGCTTGAAGACTGTCTGATGATGATCTCAAGTC 2823
 DB 849 GGGGTGAGGGGCTTTTCA-AATGATGCTTGAAGACTGTCTGATGATGATGATGATGATG 907
 QY 2824 CTTTCAACCCAGAGCTGGCCACCTCC 2852
 DB 908 CTTTCAACCCAGAGCTGGCCACCTCC 936

RESULT 13
 CF726196 705 bp mRNA linear EST 09-OCT-2003
 LOCUS CF726196
 DEFINITION U1-M-G20-c1p-h-08-0-UI.r1 NIH BMAP_G20 Mus musculus cDNA clone
 IMAGE:30606415 5', mRNA sequence.
 ACCESSION CF726196
 VERSION CF726196.1 GI:37600364
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 705)
 NIH-NCI http://nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strauberg, Ph.D.
 Email: cga@bse-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mouse1.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.

FEATURES
 Source
 Location/Qualifiers
 1..705
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30606415"
 /issue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP G20"
 /notes="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTTATGAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 23.7%; Score 688; DB 7; Length 705;
 Best Local Similarity 99.3%; Pred. No. 4.8e-151;
 Matches 688; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 31 GCGTGTGCGGTTGCGGCTGATGACTGAGCGCGGCTGAGCGGCTGCTACTGCTG 90
 DB 13 GCGTGTGCGGTTGCGGCTGATGACTGAGCGCGGCTGAGCGGCTGCTACTGCTG 72
 QY 91 CCGGACCCACTCCACTCGCGGTTCCCGCAACATGAGTGGTGGCTTACTCAGCGCC 150
 DB 73 CCGGACCCACTCCACTCGCGGTTCCCGCAACATGAGTGGTGGCTTACTCAGCGCC 132
 QY 151 CGAGCCAGGCTCCCTGCGGCTTCCGCTTGGAGAGCGCCGCGCTGGCGAGC 210
 DB 133 CGAGCCAGGCTCCCTGCGGCTTCCGCTTGGAGAGCGCCGCGCTGGCGAGC 192
 QY 211 GGCCTCATAGTGGCTTAAACCTGTGATGAAGAGAGCGGCTGAAGCGGCAACATCA 270
 DB 193 GGCCTCATAGTGGCTTAAACCTGTGATGAAGAGAGCGGCTGAAGCGGCAACATCA 252
 QY 271 AACCAACTGCGGACCGCTACGAGTTCCTGAGAGCGTGGGCAAGGGCACTACGGGA 330
 DB 253 AACCAACTGCGGACCGCTACGAGTTCCTGAGAGCGTGGGCAAGGGCACTACGGGA 312
 QY 331 AGGTGAAGAGGCAAGAGAGCTGGGCGGCTGTGGTGGCCATTAAGTCCATCAGGAAG 390
 DB 313 AGGTGAAGAGGCAAGAGAGCTGGGCGGCTGTGGTGGCCATTAAGTCCATCAGGAAG 372
 QY 391 AACCAACTGAAGAGAGAGTGTGCTGCAATACGAGAGGAGATTAAGATCAATGCTT 450
 DB 373 AACCAACTGAAGAGAGAGTGTGCTGCAATACGAGAGGAGATTAAGATCAATGCTT 432
 QY 451 CACTCAACACCCCAATCATTTGCTCATGAGTGTGTTGAGATTAAGAGCAAGATTG 510
 DB 493 TGATTGTCAAGAGATTAAGAGAGCGGAGTGTGATTAATTAATGAGAGGCGGAC 552
 QY 571 GCGTGAAGAGGCGGAGAGCGAGGATTTCTTCCAGAGATGCTGCGCTGCACTACT 630
 DB 553 GCGTGAAGAGGCGGAGAGCGAGGATTTCTTCCAGAGATGCTGCGCTGCACTACT 612
 QY 631 GCCACAGAAAGGAGTGTTCACCGAGATCTCAAGCTGAGAAACATCTTCTAGATGCA 690

Db 613 GCCACGAGAGGAGATGTTTCCACGAGATCTTCAGCTGGAGAAACATCTTCTAGATGCCA 672

QY 691 ATGGAACATCAAGATTGCTGACTTTGGCTCT 723

Db 673 NTGGAAACATCANGATTGCTGACTTTGGCTCT 705

RESULT 14
CB248251
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB248251 726 bp mRNA linear EST 15-JUL-2003
IMAGE:5718428 5', mRNA sequence.
CB248251
CB248251.1 GI:28386432
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 726)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLAN at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES
source
Seq primer: PYX-5.
Location/Qualifiers
1..726
/organism="Mus musculus"
/mol_type="mRNA"
/screen="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5718428"
/issue_type="whole brain"
/lab_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH BMAP FDO"
/note="Organ: brain; Vector: PYX-Abs; Site 1: Bcor I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lemmon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Bcor I adaptor, digested with Not I, and then
cloned directionally into PYX-Abs vector. The library tag
sequence located between the Not I site and the polyA
tail, is TGAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Query Match 23.6%; Score 683.6; DB 6; Length 726;
Best Local Similarity 98.8%; Pred. No. 5.3e-150;
Matches 720; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 922 GTAAAGGGGCTTAC-GGTGAGCGCGCCCAAGCGTCCGATGCTGTGGCTTGATCCGGTGG 980
|||||
Db 1 GTAAAGGGGCTTACGCTGAGCGCGCCCAAGCGTCCGATGCTGTGGCTTGATCCGGTGG 60
|||||
QY 981 CTGTATATGTGAACCCACCCGCTGGGCCACACTGAGAGATGTAGCCAGTCATTGGTGG 1040
|||||

Db 61 CTGTATATGTGAACCCACCCGCTGGGCCACACTGAGAGATGTAGCCAGTCATTGGTGG 120
|||||

QY 1041 GTCAACTGGGGTTACACACCGAGATCGGGGAA CAGAAAGCCCTGCTGAGGGTGGCAC 1100
|||||

Db 121 GTCAACTGGGGTTACACACCGAGATCGGGGAA CAGAAAGCCCTGCTGAGGGTGGCAC 180
|||||

QY 1101 CTTAGTGTGATCTTTGGCCGGGCTCTCATATGGCGAGCTGGTTAGTGTGCTCTCGCGCCCC 1160
|||||

Db 181 CTTAGTGTGATCTTTGGCCGGGCTCTCATATGGCGAGCTGGTTAGTGTGCTCTCGCGCCCC 240
|||||

QY 1161 CTCTGGAGATGAGAGCAAGTGTGAGCTTCTTCAAGAGACGCGGGAGGTGGA 1220
|||||

Db 241 CTCTGGAGATGAGAGCAAGTGTGAGCTTCTTCAAGAGACGCGGGAGGTGGA 300
|||||

QY 1221 AGCACTGTACTCGGCTGAGCGGCAACATTTCTTAAAGATCCCGAAGAGAAATGAC 1280
|||||

Db 301 AGCACTGTACTCGGCTGAGCGGCAACATTTCTTAAAGATCCCGAAGAGAAATGAC 360
|||||

QY 1281 ATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGAGATCTCTTCTGCTGCGCAAGAC 1340
|||||

Db 361 ATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGAGATCTCTTCTGCTGCGCAAGAC 420
|||||

QY 1341 AGCTTAAGCTTCGAAAGGCAATTCAGAAAAAGCTCTACCTCTGAGGGGGGTA 1400
|||||

Db 421 AGCTTAAGCTTCGAAAGGCAATTCAGAAAAAGCTCTACCTCTGAGGGGGGTA 480
|||||

QY 1401 CAGAGAGACCTCAGAGAACTCAGACCGGTGCTGATATCTCAGAGGAGCTGTCCCTGT 1460
|||||

Db 481 CAGAGAGACCTCAGAGAACTCAGACCGGTGCTGATATCTCAGAGGAGCTGTCCCTGT 540
|||||

QY 1461 GTATCCCTGCTCCAGAGAAAGCAATCTTAAAGTCTGACAGCGTGAATCTGGTTAC 1520
|||||

Db 541 GTATCCCTGCTCCAGAGAAAGCAATCTTAAAGTCTGACAGCGTGAATCTGGTTAC 600
|||||

QY 1521 TACTCCCTCAGAGGCGAGGAGTCTGGGGAACTCTTAAAGCGCATGTATGTGTTG 1580
|||||

Db 601 TACTCCCTCAGAGGCGAGGAGTCTGGGGAACTCTTAAAGCGCATGTATGTGTTG 660
|||||

QY 1581 AGTGGGAGCCCGGTGAGAGCAAGTCTCAGAGGCTTCTGCTTCCACCGCAAG 1640
|||||

Db 661 AGGAGGAGCA-CCCGTGGAGAGAA--TCTACAGAGCTTCAGGGGCTCTCTCCACCGCAAG 717
|||||

QY 1641 GGCATTCTC 1649
|||||

Db 718 GGCATTCTC 726
|||||

RESULT 15
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CO424322 672 bp mRNA linear EST 06-JUL-2004
IMAGE:30665189 5', mRNA sequence.
CO424322
CO424322.1 GI:49670481
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 672)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. .672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:3065389"
/tissue_type="whole eye"
/dev_stage="newborn (1, 5, 15 days)"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="NIH BMAP_H00"
/note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AATAATTACG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 23.2%; Score 672; DB 7; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.8e-147;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 913 AGCAATCAGTAAGGGGCTTACCGTAGCGGCCCAAGCGTCCGATGCGTGGCCCTGA 972
Db 1 AGCAATCAGTAAGGGGCTTACCGTAGCGGCCCAAGCGTCCGATGCGTGGCCCTGA 60
QY 973 TCCGGTGGCTGTATGTGTAACCCCAACCCCTCGGGCCACACTGAGAGATGACCAATC 1032
Db 61 TCCGGTGGCTGTATGTGTAACCCCAACCCCTCGGGCCACACTGAGAGATGACCAATC 120
QY 1033 ATTGGTGGGTCACTGAGGCTTACCAACCGGAGTGGGGAAACAGAAAGCCCTGCGTAGG 1092
Db 121 ATTGGTGGGTCACTGAGGCTTACCAACCGGAGTGGGGAAACAGAAAGCCCTGCGTAGG 180
QY 1093 GTGGGCACCCCTAGTGTGTAACCTTGGCCGGGCTCCATGCGGACTGGTTACGTGCTCCT 1152
Db 181 GTGGGCACCCCTAGTGTGTAACCTTGGCCGGGCTCCATGCGGACTGGTTACGTGCTCCT 240
QY 1153 CGCGCCCTCTCTGAGAGATGAGCAAGGTGTGCAAGCTTTCAGACAGACGTCGCGG 1212
Db 241 CGCGCCCTCTCTGAGAGATGAGCAAGGTGTGCAAGCTTTCAGACAGACGTCGCGG 300
QY 1213 GAGGTGGAAGCACTGTACTGTGGGCTGAGCGGCAACATTTCTTAAGAAATCCGAAAGG 1272
Db 301 GAGGTGGAAGCACTGTACTGTGGGCTGAGCGGCAACATTTCTTAAGAAATCCGAAAGG 360
QY 1273 AGAATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTTTCTGCGCCTG 1332
Db 361 AGAATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTTTCTGCGCCTG 420
QY 1333 GCAAGAGAGAGCTTAAGCTTCCGAAAGGCAATTTCAAGAAAGAGCTCTACCTGCTGAG 1392
Db 421 GCAAGAGAGAGCTTAAGCTTCCGAAAGGCAATTTCAAGAAAGAGCTCTACCTGCTGAG 480
QY 1393 GGAAGGTACAGAGAGAGCCCTCAGAACTCAGACCGGTGCTGATATCTCAGGGGACCTG 1452
Db 481 GGAAGGTACAGAGAGAGCCCTCAGAACTCAGACCGGTGCTGATATCTCAGGGGACCTG 540
QY 1453 TCCCTGCTGTATCCCTGCTCCCAAGGAAAGGATCTTAAGAAGTCTGACAGGTGAAT 1512
Db 541 TCCCTGCTGTATCCCTGCTCCCAAGGAAAGGATCTTAAGAAGTCTGACAGGTGAAT 600
```

```
QY 1513 CTGTTACTACTCTCTCTCCAGAGCCAGCGAGTGGGGAATCTTGAAGCCAGTAGTG 1572
Db 601 CTGTTACTACTCTCTCTCCAGAGCCAGCGAGTGGGGAATCTTGAAGCCAGTAGTG 660
QY 1573 TGTTTGAGTG 1584
Db 661 TGTTTGAGTG 672
```

Search completed: May 16, 2005, 03:15:24
Job time : 9317.58 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 15, 2005, 22:46:12 ; Search time 1655.86 Seconds
(without alignments)
10726.042 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902
Sequence: 1 cactagtgcagtcacgaagt.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodaca/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodaca/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodaca/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodaca/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodaca/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodaca/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodaca/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodaca/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodaca/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodaca/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodaca/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodaca/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodaca/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodaca/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodaca/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodaca/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodaca/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodaca/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodaca/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodaca/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodaca/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodaca/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	US-10-355-975-4	Sequence 4, Appl
2	2863.4	98.7	3073	US-10-322-281-22	Sequence 22, Appl
3	2066.8	71.2	2929	US-10-343-514-1	Sequence 1, Appl
4	2022.2	69.7	2027	US-10-343-514-87	Sequence 87, Appl
5	1938.8	66.8	37278	US-10-322-281-21	Sequence 21, Appl
6	1771.4	61.0	2026	US-10-343-514-27	Sequence 27, Appl
7	1379	47.5	3353	US-09-963-159-1	Sequence 1, Appl
8	1379	47.5	3353	US-10-423-543-43	Sequence 43, Appl
9	1375.8	47.4	3404	US-10-322-281-25	Sequence 25, Appl
10	1374.2	47.4	3443	US-10-370-715B-639	Sequence 639, Appl
11	1374.2	47.4	3463	US-10-618-941-11	Sequence 11, Appl

12	1372.2	47.3	2501	17	US-10-302-172-215	Sequence 215, Appl
13	1370.6	47.2	3360	17	US-10-311-034-38	Sequence 38, Appl
14	1360.2	46.9	1884	9	US-09-963-159-3	Sequence 3, Appl
15	1360.2	46.9	1884	17	US-10-423-543-45	Sequence 13, Appl
16	1071.4	36.9	1186	18	US-10-343-514-13	Sequence 13, Appl
17	943.4	32.5	1186	18	US-10-343-514-70	Sequence 70, Appl
18	849.2	29.3	2616	18	US-10-363-829-151	Sequence 250, Appl
19	849.2	29.3	2619	18	US-10-363-829-250	Sequence 24, Appl
20	732	25.2	39699	18	US-10-322-281-24	Sequence 32, Appl
21	675.4	23.3	741	18	US-10-343-514-32	Sequence 91, Appl
22	648	22.3	707	18	US-10-343-514-91	Sequence 31, Appl
23	648	22.3	712	18	US-10-343-514-31	Sequence 93, Appl
24	644.2	22.2	661	18	US-10-343-514-93	Sequence 92, Appl
25	643	22.2	734	18	US-10-343-514-92	Sequence 97, Appl
26	583.2	20.1	588	18	US-10-343-514-97	Sequence 33, Appl
27	557.4	19.2	660	18	US-10-343-514-33	Sequence 95, Appl
28	534	18.4	578	18	US-10-343-514-95	Sequence 96, Appl
29	515	17.7	548	18	US-10-343-514-96	Sequence 94, Appl
30	513	17.7	521	18	US-10-343-514-94	Sequence 35, Appl
31	511.6	17.6	579	18	US-10-343-514-35	Sequence 1, Appl
32	507.6	17.5	6828	9	US-09-780-949-1	Sequence 81, Appl
33	507.6	17.5	6828	16	US-10-354-358-81	Sequence 1541, Ap
34	507.6	17.5	6828	17	US-10-172-118-1541	Sequence 2552, Ap
35	507.6	17.5	6828	17	US-10-342-887-1541	Sequence 6703, Ap
36	507.6	17.5	6828	18	US-10-723-860-2552	Sequence 3, Appl
37	507.6	17.5	6854	18	US-10-723-860-6703	Sequence 37, Appl
38	506.4	17.5	1986	9	US-09-780-949-5	Sequence 36, Appl
39	506.4	17.5	2884	9	US-10-343-514-37	Sequence 34, Appl
40	493.6	17.0	585	18	US-10-343-514-36	Sequence 10, Appl
41	483.8	16.7	548	18	US-10-343-514-34	Sequence 14, Appl
42	474.6	16.4	521	18	US-10-343-514-10	Sequence 9, Appl
43	383.4	13.2	439	18	US-10-343-514-14	
44	381.4	13.1	436	18	US-10-343-514-14	
45	365.6	12.6	421	18	US-10-343-514-9	

ALIGNMENTS

RESULT 1
US-10-355-975-4
; Sequence 4, Application US/10355975
; Publication No. US20030162277A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Vitca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MORINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/10/355, 975
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US/09/579, 664B
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-355-975-4

Query Match 100.0%; Score 2902; DB 16; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 CACTAGTGATCCAAAGATTCCGACGAGCGTGTCTGGGTGGGCTGTGACTCTGAG 60
1 CACTAGTGATCCAAAGATTCCGACGAGCGTGTCTGGGTGGGCTGTGACTCTGAG 60
CCCGGGCTCAGCGCGGCTGTCTGCTGCTCCGACCACTCCACTCTGCGGTCCGCA 120

Db 61 CCCGCGCTCAGCGCGCTCTACTGCTGCCGACCACTCACTCGCGCTCCCGCA 120
 QY 121 CCATGAGTGGTGGCTTAACTCCAGCGCCCGAGACGAGCTCCTCGGCTCCGCTCGG 180
 Db 121 CCATGAGTGGTGGCTTAACTCCAGCGCCCGAGACGAGCTCCTCGGCTCCGCTCGG 180
 QY 181 CCTCGAGAGCGCCCGCGCTGGCGGACGCGCTCATCAAGTGGCTTAACTCGTGA 240
 Db 181 CCTCGAGAGCGCCCGCGCTGGCGGACGCGCTCATCAAGTGGCTTAACTCGTGA 240
 QY 241 AGAAGCAGCGGTGAGCGGACCATCAAAACAACCTGCGGACCGCTTACAGATTC 300
 Db 241 AGAAGCAGCGGTGAGCGGACCATCAAAACAACCTGCGGACCGCTTACAGATTC 300
 QY 301 TGAAGAGCTGGGGAAGGACCTTACGCGGAAGGTGAAGAGGACGAGAGCTGGG 360
 Db 301 TGAAGAGCTGGGGAAGGACCTTACGCGGAAGGTGAAGAGGACGAGAGCTGGG 360
 QY 361 GTCGTGGGCTCATCAATCCATCAAGAAAGCAAAATCAAAATGAGAGAGATTCG 420
 Db 361 GTCGTGGGCTCATCAATCCATCAAGAAAGCAAAATCAAAATGAGAGAGATTCG 420
 QY 421 ACATACGAGGAGGATTTGAGATCATGTCTTCACTCAACACACCCCACTATTC 480
 Db 421 ACATACGAGGAGGATTTGAGATCATGTCTTCACTCAACACACCCCACTATTC 480
 QY 481 ATGAAGTGTGGAATATGAGCAAGATTTGATGATGATGATGATGATGATGATG 540
 Db 481 ATGAAGTGTGGAATATGAGCAAGATTTGATGATGATGATGATGATGATGATG 540
 QY 541 ATCTGATGATTAATCATGATGAGCGGACCGCTGATGAGCGGACCGGACGATTC 600
 Db 541 ATCTGATGATTAATCATGATGAGCGGACCGCTGATGAGCGGACCGGACGATTC 600
 QY 601 TCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db 601 TCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 661 TCAAGCTGGAAGAAATCTCTTCTGATGATGATGATGATGATGATGATGATG 720
 Db 661 TCAAGCTGGAAGAAATCTCTTCTGATGATGATGATGATGATGATGATGATG 720
 QY 721 TCTTCAACCTGATCAAAAGGCAAGTTCCTCCAGAGCTTCTGAGAGGCTCTCA 780
 Db 721 TCTTCAACCTGATCAAAAGGCAAGTTCCTCCAGAGCTTCTGAGAGGCTCTCA 780
 QY 781 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 Db 781 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 841 TGGGCGCTTCTCTGATCAATCTGATGATGATGATGATGATGATGATGATG 900
 Db 841 TGGGCGCTTCTCTGATCAATCTGATGATGATGATGATGATGATGATGATG 900
 QY 901 AAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 Db 901 AAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 QY 961 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 Db 961 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1021 ATGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 Db 1021 ATGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1081 CCTGCGTGAAGGAGGACCTAGTGTGATGATGATGATGATGATGATGATGATG 1140
 Db 1081 CCTGCGTGAAGGAGGACCTAGTGTGATGATGATGATGATGATGATGATGATG 1140
 QY 1141 TACGTGCTCTCGCGCTCCCTCTGAGAGATGAGCAAGTGTGAGCTTCTTCAAG 1200
 Db 1141 TACGTGCTCTCGCGCTCCCTCTGAGAGATGAGCAAGTGTGAGCTTCTTCAAG 1200

QY 1201 AGCACGTCGCGGAGAGTGAAGACATGATCTGAGCTGAGAGCGGCAATTTCTTTA 1260
 Db 1201 AGCACGTCGCGGAGAGTGAAGACATGATCTGAGCTGAGAGCGGCAATTTCTTTA 1260
 QY 1261 AGTCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 Db 1261 AGTCCGAAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 QY 1321 CTCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTTCTTCAAGAA 1380
 Db 1321 CTCTGCGCTTGGCAAGAGAGCTTAAAGCTTCCGAAAGGCAATTTCTTCAAGAA 1380
 QY 1381 CTACTGCTCAAGGAGAGTCAAGAGAGAGCTTCAAGAACTTCAAGAGCTGATG 1440
 Db 1381 CTACTGCTCAAGGAGAGTCAAGAGAGAGCTTCAAGAACTTCAAGAGCTGATG 1440
 QY 1441 CAGGAGAGCTGTCCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1500
 Db 1441 CAGGAGAGCTGTCCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1500
 QY 1501 GACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
 Db 1501 GACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
 QY 1561 ACGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 Db 1561 ACGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 QY 1621 GGTCTCTCTCCACCGGCAAGGAGCTTCAAACTCAATGAGCAAGTTCCTCCG 1680
 Db 1621 GGTCTCTCTCCACCGGCAAGGAGCTTCAAACTCAATGAGCAAGTTCCTCCG 1680
 QY 1681 TAGAAGCACTACCCCTAGACCTTGGCTCCCTGACCAATGAGCTTCCCATCT 1740
 Db 1681 TAGAAGCACTACCCCTAGACCTTGGCTCCCTGACCAATGAGCTTCCCATCT 1740
 QY 1741 CAGCCGCGCCAGCGCTCTGAGGCTGATGATGATGATGATGATGATGATGATG 1800
 Db 1741 CAGCCGCGCCAGCGCTCTGAGGCTGATGATGATGATGATGATGATGATGATG 1800
 QY 1801 CTTTGAACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
 Db 1801 CTTTGAACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
 QY 1861 CTGTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 Db 1861 CTGTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 QY 1921 AGGAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
 Db 1921 AGGAATCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
 QY 1981 ACAAGCAAGGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 2040
 Db 1981 ACAAGCAAGGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 2040
 QY 2041 ATGGGATGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTT 2100
 Db 2041 ATGGGATGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTT 2100
 QY 2101 GAGTACATCAAGGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAG 2160
 Db 2101 GAGTACATCAAGGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAG 2160
 QY 2161 ATTGAATGGAAGAGATGGAAGCTTGTGCTGAGAGCTTGAAGAGCTTGAAGAGAT 2220
 Db 2161 ATTGAATGGAAGAGATGGAAGCTTGTGCTGAGAGCTTGAAGAGCTTGAAGAGAT 2220
 QY 2221 GGTGCTACTCTCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 2280
 Db 2221 GGTGCTACTCTCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 2280

Db	1141	TGAGCCAGGTTGCACTTCTCAAGCAGCAGTGCCGGAGGTGGAACACTGTACC	1200
Qy	1232	TGGGCTGAGCGGGAACATTTCTTAAGAAAGTCCCGAAAGAAATGACATGCTCAAA	1291
Db	1201	TGGCTGAGCGGGAACATTTCTTAAGAAAGTCCCGAAAGAAATGACATGCTCAAA	1260
Qy	1292	TCTGCAAGTACCCGGCTGAGATACCTCTTCTGCGCTGGCAAGAGCCTTAAGCT	1351
Db	1261	TCTGCAAGTACCCGGCTGAGATACCTCTTCTGCGCTGGCAAGAGCCTTAAGCT	1320
Qy	1352	TCCGAAGGCAATTCAGAAAAAGTCTTAAGTCTGTCAGAGGAGGTACAGAGAAC	1411
Db	1321	TCCGAAGGCAATTCAGAAAAAGTCTTAAGTCTGTCAGAGGAGGTACAGAGAAC	1380
Qy	1412	TCAGGAATCAGACCGGCTGATATCTCCAGGGAGCCTTCTGCTGTATCTCTGCT	1471
Db	1381	TCAGGAATCAGACCGGCTGATATCTCCAGGGAGCCTTCTGCTGTATCTCTGCT	1440
Qy	1472	CCCAAGGAAGGCAATTCCTTAAGAGTCTGCAAGAGGATCTGATTAATCTCTCTCC	1531
Db	1441	CCCAAGGAAGGCAATTCCTTAAGAGTCTGCAAGAGGATCTGATTAATCTCTCTCC	1500
Qy	1532	AGAGCCAGCGAGTCTGAGGAACTTTAGACGCGAGATGATGTTGTGAGTGGGACC	1591
Db	1501	AGAGCCAGCGAGTCTGAGGAACTTTAGACGCGAGATGATGTTGTGAGTGGGACC	1560
Qy	1592	CGTGAAGCAAGTCTCCAGGCTTCAAGGCTTCTCTCCACCGCAAGGCAATTCCTCA	1651
Db	1561	CGTGAAGCAAGTCTCCAGGCTTCAAGGCTTCTCTCCACCGCAAGGCAATTCCTCA	1620
Qy	1652	ACTCAATGGCAAGTCTCCGCAAGGCTTGAAGGCACTACCCCTGACCTTTGAGCTC	1711
Db	1621	ACTCAATGGCAAGTCTCCGCAAGGCTTGAAGGCACTACCCCTGACCTTTGAGCTC	1680
Qy	1712	CTGGAACCAACTGAGCTCTCTCCATCTCTGACGCGGCGCCAGCGCTTCAAGGAGCT	1771
Db	1681	CTGGAACCAACTGAGCTCTCTCCATCTCTGACGCGGCGCCAGCGCTTCAAGGAGCT	1740
Qy	1772	GAGTGAAGGAGCACTCTGCTCTCCGAGCTCTTGAACCAATGGAATCTGCTGAAGCT	1831
Db	1741	GAGTGAAGGAGCACTCTGCTCTCCGAGCTCTTGAACCAATGGAATCTGCTGAAGCT	1800
Qy	1832	TCCGGAACCCCACTGAGGAGCTGTGTCTGTGACCAACTGAGGAGGCTTGAAGACC	1891
Db	1801	TCCGGAACCCCACTGAGGAGCTGTGTCTGTGACCAACTGAGGAGGCTTGAAGACC	1860
Qy	1892	TCCCTCAGAAAGTCTGAAGCGATGAGGAGCAAGAACTTGGGAGATGCTCTTCTCT	1951
Db	1861	TCCCTCAGAAAGTCTGAAGCGATGAGGAGCAAGAACTTGGGAGATGCTCTTCTCT	1920
Qy	1952	GAGAGCTGCGCAAGGATGATGAGGCTTACAGCAAGCCCTTGAAGATCTGCTCAAGCT	2011
Db	1921	GAGAGCTGCGCAAGGATGATGAGGCTTACAGCAAGCCCTTGAAGATCTGCTCAAGCT	1980
Qy	2012	CAGCTGAGGAAGGAGATGAGGCTTGTATGAGGAGTCTGAGAGGAGGTTGCAAGG	2071
Db	1981	CAGCTGAGGAAGGAGATGAGGCTTGTATGAGGAGTCTGAGAGGAGGTTGCAAGG	2040
Qy	2072	AACCTGAGGATGATCTCAAGTAAATGAGTACATCAAGGAGCTTACGCTGCAAGCT	2131
Db	2041	AACCTGAGGATGATCTCAAGTAAATGAGTACATCAAGGAGCTTACGCTGCAAGCT	2100
Qy	2132	GACTGAACCTGAAAGATGAGAAATGCAATGATGAGAAAGATGAGAACTTGTCT	2191
Db	2101	GACTGAACCTGAAAGATGAGAAATGCAATGATGAGAAAGATGAGAACTTGTCT	2160
Qy	2192	GCCGAGTGTATAGTGGGAGGCTGAGGAGTCTACCTTGTGCAATGAGTCA	2251
Db	2161	GCCGAGTGTATAGTGGGAGGCTGAGGAGTCTACCTTGTGCAATGAGTCA	2220
Qy	2252	CCCATGACATTTCCCAACCTGTTCTCTGCTGCACTTCAATAGTTCTGTTCCATC	2311
Db	2221	CCCATGACATTTCCCAACCTGTTCTCTGCTGCACTTCAATAGTTCTGTTCCATC	2280

Qy	2312	AAACCAAGGAGTTAAGAACCTGACTTCTGGAAGTATGTATGATGACATGCTTATTT	2371
Db	2281	AAACCAAGGAGTTAAGAACCTGACTTCTGGAAGTATGTATGATGACATGCTTATTT	2340
Qy	2372	AGAGAGGAACAGGCTGTGGTTTCCATCTGCTGCTGTCATCTCAAGACCTGGGAAG	2431
Db	2341	AGAGAGGAACAGGCTGTGGTTTCCATCTGCTGCTGTCATCTCAAGACCTGGGAAG	2400
Qy	2432	ACTGGAACGAGCTGTTGACTTCAATCTCAAGGGAACAGATGCCCCCTGAGACCTTAA	2491
Db	2401	ACTGGAACGAGCTGTTGACTTCAATCTCAAGGGAACAGATGCCCCCTGAGACCTTAA	2460
Qy	2492	ATCTCAGAGCTTGAACCTTGAAGCTTCTCTAGTACCCAGATGATGATGATGCTCTGT	2551
Db	2461	ATCTCAGAGCTTGAACCTTGAAGCTTCTCTAGTACCCAGATGATGATGATGCTCTGT	2520
Qy	2552	TTCTCAGGCAAGGAGCACTAGATGTCGACTTATTTATTTTGTGATCTCACTT	2611
Db	2521	TTCTCAGGCAAGGAGCACTAGATGTCGACTTATTTATTTTGTGATCTCACTT	2580
Qy	2612	CTGTTTTTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	2671
Db	2581	CTGTTTTTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	2640
Qy	2672	AATTAATGTAATGCTGTGTTCTGGGAACTCCACTGTCGCACTGAAGTTATGTAAGAG	2731
Db	2641	AATTAATGTAATGCTGTGTTCTGGGAACTCCACTGTCGCACTGAAGTTATGTAAGAG	2700
Qy	2732	AGGATTTGGCAATGATGCTCTCTATTTCAAGGAGGAGTGGGAGGCTTTTCAATGATG	2791
Db	2701	AGGATTTGGCAATGATGCTCTCTATTTCAAGGAGGAGTGGGAGGCTTTTCAATGATG	2760
Qy	2792	TCTTGAAGCACTGTCTGATTAAGTCTTCAAGGAGGAGTGGGAGGCTTTTCAATGATG	2851
Db	2761	TCTTGAAGCACTGTCTGATTAAGTCTTCAAGGAGGAGTGGGAGGCTTTTCAATGATG	2820
Qy	2852	CTCATCTTCACTGTGTGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2896
Db	2821	CTCATCTTCACTGTGTGCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2865

RESULT 3
 US-10-343-514-1
 ; Sequence 1, Application US/10343514
 ; Publication No. US20040132025A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUCKER, Daniel J.
 ; APPLICANT: ROSEN, Cheryl P.
 ; APPLICANT: LEBEYRE, Diana L.
 ; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
 ; FILE REFERENCE: DPA-DRUC2/PCT
 ; CURRENT APPLICATION NUMBER: US/10/343,514
 ; CURRENT FILING DATE: 2003-01-31
 ; PRIOR APPLICATION NUMBER: PCT/CA01/01109
 ; PRIOR FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: US 60/222,650
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/274,613
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: CA 2,340,780
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2929
 ; TYPE: DNA
 ; ORGANISM: RAT
 ; US-10-343-514-1
 Query Match 71.2%; Score 2066.8; DB 18; Length 2929;
 Best Local Similarity 86.4%; Pred. No. 0;
 Matches 2488; Conservative 0; Mismatches 277; Indels 115; Gaps 14;

QY 49 GTGACCTCTGAGCCGCGGCTCAGCGCGCGCTGTGCTGCTGCCGACCCACTCCACCTC 108
DB 9 GTGACCTCTGAGCTGCGGCTCTCCGCGCGCTGCTGCTGCCGACCCCTCCGCTC 68
QY 109 GCGGTCCTCCGACCATGTGAGTCCGCTTATCTCTCAGCGCCCGAGCCAGGCTCTCTCG 168
DB 69 GCGCTCCCGACCATGTGAGTCCGCTTATCTCTCAGCGCCCGAGCCAGGCTCTCTCG 128
QY 169 CCTCGCGCTGAGCTCCGAGAGCGCCGCGCTGCGGAGCGGGCTCATCAAGTCGCTTA 228
DB 129 CTTCCGCTGAGCATCGAGAGCGCCGCGCTGCGGAGCGGGCTCATCAAGTCGCTTA 188
QY 229 AACCTCTGATGAGAGAGCGCGGTGAGCGGACCATCAACCAACCACTCTCGGACCC 288
DB 189 AACCTCTGATGAGAGAGCGCGGTGAGCGGACCATCAACCAACCACTCTCGGACCC 248
QY 289 GCTACGAGTCTCTGAGAGCGCTGCGGACCATCAACGAGGAGGTGAGAGAGCGAG 348
DB 249 GCTACGAGTCTCTGAGAGCGCTGCGGACCATCAACGAGGAGGTGAGAGAGCGAG 308
QY 349 AGAGCTGAGGAGCGCTGCTGAGCATCAAGTCATAGGAGAGAGAGAGAGAGAGAG 408
DB 309 AGAGCTGAGAGCGCTGCTGAGCATCAAGTCATAGGAGAGAGAGAGAGAGAGAGAG 368
QY 409 AGGATCTGCTGACATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
DB 369 AGGATCTGCTGACATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
QY 469 TCATTCGATTCATGAG 528
DB 429 TCATTCGATTCATGAG 488
QY 529 CCGAGCGAG 588
DB 489 CCGAGCGAG 548
QY 589 CCGAGCGAG 648
DB 549 CCGAGCGAG 608
QY 649 TTCAACGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
DB 609 TTCAACGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
QY 709 CTGACTTTGGGCTCTCCAGCTGTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
DB 669 CTGACTTTGGGCTCTCCAGCTGTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 728
QY 769 GCGCTCTCTACGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
DB 729 GCGCTCTCTACGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 788
QY 829 ACAAGCTGTCTCTGAG 888
DB 789 ACAAGCTGTCTCTGAG 848
QY 889 GGCAGAGATCATAAACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
DB 849 GGCAGAGATCATAAACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 908
QY 949 AGCGCTCGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
DB 909 AGCGCTCGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968
QY 1009 CCACACTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
DB 969 CCACACTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1028
QY 1069 GGGAGACGAG 1128
DB 1029 GGGAGACGAG 1088

QY 1129 TGGCGAGCTGAGTACGTGCTCTCTCGCGCCCTCTCTCTGAGAGATGAGAGCCAGGTGTC 1188
DB 1089 TGGCGAGCTGAGTACGTGCTCTCTCTCGCGCCCTCTCTCTGAGAGATGAGAGCCAGGTGTC 1148
QY 1189 GCTTCTTCAAG 1248
DB 1149 GCTTCTTCAAG 1208
QY 1249 ATTCTCTTAAAGATGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
DB 1209 ATTCTCTTAAAGATGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
QY 1309 CTGAGATACCT 1368
DB 1269 CTGAGATACCT 1328
QY 1369 AGAGAAAGCT 1428
DB 1329 AGAGAAAGCT 1388
QY 1429 TGCTGATACCT 1488
DB 1389 TGCTGATACCT 1448
QY 1489 TTAAGAGTCTCGACAGCGTGAATCTGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1548
DB 1449 TTAAGAGTCTCGACAGCGTGAATCTGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1508
QY 1549 GGGAGCTCTTAAAG 1608
DB 1509 GGGAGCTCTTAAAG 1568
QY 1609 CACAGGCTTCAAG 1668
DB 1569 CACAGGCTTCAAG 1628
QY 1669 CCGGACAGAGCTTGAAG 1728
DB 1626 CCGGACAGAGCTTGAAG 1685
QY 1729 CTTCCATCTCTGAG 1788
DB 1686 CTTCCATCTCTGAG 1745
QY 1789 TGTCTCTCGAGTCTTGAAG 1848
DB 1746 TGTCTCTCGAGTCTTGAAG 1805
QY 1849 GGGAGCTGTGTCTGTGAG 1908
DB 1806 GGGAGCTGTGTCTGTGAG 1865
QY 1909 AGGAGTGTGTGAG 1968
DB 1866 AGGAGTGTGTGAG 1925
QY 1969 TGAAG 2028
DB 1926 TGAAG 1985
QY 2029 TGTGCTCTCTAGT-ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2087
DB 1986 TGTGCTCTCTAGT-ATGAG 2044
QY 2088 CTTCAAGTGAATGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2143
DB 2045 CTTCAAGTGAATGAGATCAAG 2104
QY 2144 AAGATGAG 2203
DB 2105 AAGATGAG 2164
QY 2204 TAGTGGAGT--GAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2260

Db 2165 TGAGCGGACAGAGACGAAAGGTGCTACCTCTTATGC--TGAGTGCTACCCATGGCA 2222
Qy 2261 TTTCCCAACCTGTTCTCTGGTGCAAC-----TTCACTAAGTTTCTGTTCCATCA 2312
Db 2223 TCTCCC-CCCTGCTCTCTGCGCAGGTGACGGGTGACCACTAAGTTCCTGTTGCACTG 2281
Qy 2213 ACCACGAGGTAGAACCTCTGACTTCCGAGGAGTAATGTAGTGAATGTCATTATTTA 2312
Db 2282 ACCACGAGGTAGAACCTCTGACATCCCTGGAATATGTAGAACCACTGCTTATTTA 2341
Qy 2373 GAGAGAAACAGCTCTGTTTTCATCTCTGCTGTCATCTCAAGAAGCTGGAAGA 2432
Db 2342 AAGAGAAACAGCTCTGTTTTCATCTCTGCTGTCATCTCAAGAAGCTGGAAGA 2401
Qy 2433 CTGGAGCCGCTGTTTGAATCTCAAGGGAGCAAGTGCCTGAGCCCATCTTGA 2492
Db 2402 CTC-ACTGCTGTTCACTTCATCTCAAGGGAGC----- 2434
Qy 2493 TCTCAGAGCTTGAACCTTGAAGCTGTTCTAGTACCAGATGAGTGAATGAT- GCTGCT 2551
Db 2435 -CTCAGAGACTGAGCTTGAAGCTGTTCTAGTACCAGATGAGTGAATGATGCTGT 2493
Qy 2552 TTCTCAGGCCAAGCGGAGCTAGAAATGTGCTGACTTATTTTGTGATTCCTCACTT 2611
Db 2494 TTCTCAGGCCAAGCGGAGCCAGAAATGTGCTGACTTATTTA--TTTTGTGATTCCTCACTT 2551
Qy 2612 CTGTTTTTGGTT----- 2624
Db 2552 CTGTTTCTGTTTCTGTTTGTGTTGTGTTTGTGTTGTTGTTGTTGTTTCTGTTT 2611
Qy 2625 -----TTGTTGTTGTTGTTGTTTGTGTTTGAAGTAATTTGCTGCTTCAATATG 2618
Db 2612 TGTTTGTGTTGTTGTTGTTGTTGTTTGAAGTAATTTGCTGCTTCAATATG 2611
Qy 2679 TGAATGCTGTTGTTGAGGAACTCACCTGTCCTGAGTGAATTTATGACAGAGATTT 2738
Db 2672 TGAATGCTGTTGTTGAGGAAAGCCACGTGTGATGAGTGTGACAGAGATTT 2731
Qy 2739 TGGCAATGATGTCCTCTATTCAGAGGGGGGTGGGGCGTTTCAAAATGATGCTTGA 2798
Db 2732 TGGCAATGATGTCCTCTATTCAGAGGGGGGTGGGGCGTTTCAAAATGATGCTTGA 2783
Qy 2799 CACTGTGATGATGCTCAGTCCCTTCAACCAAGGCTGGCCACCTCCCTCATCT 2858
Db 2784 CACTGTGATGATGCTCAGTCCCTTCAACCAAGGCTGGCCACCTCCCTCATCT 2843

RESULT 4
US-10-343-514-87
; Sequence 87, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEBEYRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUCZ/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87
; LENGTH: 2027
; TYPE: DNA

; ORGANISM: MOUSE
US-10-343-514-87
Query Match 69.7%; Score 2022.2; DB 10; Length 2027;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2024; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 49 GTGACCTCTAGGCGCGGCTCAAGCGGCGGCTGATCTGTCGCCGCGGACCACTCACTC 108
Db 1 GTGACCTCTAGGCGCGGCTCAAGCGGCGGCTGATCTGTCGCCGCGGACCACTCACTC 60
Qy 109 GCGGTCCCGCCACCATGAGTCCGTGCTTACTTCCAGCCCGAGCGGCTCCCTCG 168
Db 61 GCGGTCCCGCCACCATGAGTCCGTGCTTACTTCCAGCCCGAGCGGCTCCCTCG 120
Qy 169 CTTCCGCTGCTGCTGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTGCCTA 228
Db 121 CTTCCGCTGCTGCTGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTGCCTA 180
Qy 229 AACCTGATGAG 288
Db 181 AACCTGATGAG 240
Qy 289 GCTACGAGTTCCTGAG 348
Db 241 GCTACGAGTTCCTGAG 300
Qy 349 AGAGCTCGGGCGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
Db 301 AGAGCTCGGGCGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 409 AGAGCTCGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
Db 361 AGAGCTCGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 469 TCAATGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
Db 421 TCAATGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 529 CAGGCGAGGCGATCTGATGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 588
Db 481 CAGGCGAGGCGATCTGATGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 589 CAGGCGATTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
Db 541 CAGGCGATTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 649 TTACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
Db 601 TTACCGAGATCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 709 CTGACTTTGGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
Db 661 CTGACTTTGGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 769 GCCCTCTTACGCTGCTGAGATGATCAAGGAGAGAGAGAGAGAGAGAGAGAG 828
Db 721 GCCCTCTTACGCTGCTGAGATGATCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 829 ACAGCTGCTCTGAGGAGGCTCTCTGATCAATCTGAGAGAGAGAGAGAGAGAGAG 888
Db 781 ACAGCTGCTCTGAGGAGGCTCTCTGATCAATCTGAGAGAGAGAGAGAGAGAGAG 840
Qy 889 GAGAGATCATTAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
Db 841 GAGAGATCATTAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 949 AGCGTCCGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
Db 901 AGCGTCCGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 1009 CCAACTGAGAGATGAG 1068

Db	961	CCACA	CTGAGAGATGTATAC	CAGTACATTTGTGGGTCAAC	CTGGGGTTTACAC	CCGAGTCG	1020
Qy	1069	GGGAA	CAGGAAAGCCCTG	GTGAGGGTGGGCA	CCCTAAGTGTGAC	CTTTTGCCGGGCTTCCA	11228
Db	1021	GGGAA	CAGAAAGCCCTG	GTGAGGGTGGGCA	CCCTAAGTGTGAC	CTTTTGCCGGGCTTCCA	1080
Qy	1129	TGGCGGA	CTGGTATAGTGC	CTCTCGCCCTCCTT	CGAAGATAGAAC	CAAGGTGTCA	1188
Db	1081	TGGCGGA	CTGGTATAGTGC	CTCTCGCCCTCCTT	CGAAGATAGAAC	CAAGGTGTCA	11404
Qy	1189	GCTTCTT	CAAGCAAGCAGTGC	CCGGAGGTGAAAGCATGTAC	CTGGGCTGGAGCGGCAAC	1248	
Db	1141	GCTTCTT	CAAGCAAGCAGTGC	CCGGAGGTGAAAGCATGTAC	CTGGGCTGGAGCGGCAAC	1200	
Qy	1249	ATTTCCTT	TAAGAAATGCCGAAAGAAATGAC	ATGGCTCAAAATCTGAAAGTGA	CCCGG	1308	
Db	1201	ATTTCCTT	TAAGAAATGCCGAAAGAAATGAC	ATGGCTCAAAATCTGAAAGTGA	CCCGG	1260	
Qy	1309	CTGAGS	ATACCTCTTCTG	CCCTGGCAAGACAGCCTTAAGCTTCCGAAAGGCA	ATTCTCA	1368	
Db	1261	CTGAGS	ATACCTCTTCTG	CCCTGGCAAGACAGCCTTAAGCTTCCGAAAGGCA	ATTCTCA	1320	
Qy	1369	AGAAAA	AGTCTCTA	CTCTGTCAAGGGAGTACAGAGAGAC	CTTCAGGA	ACTCAGACCG	1428
Db	1321	AGAAAA	AGTCTCTA	CTCTGTCAAGGGAGTACAGAGAGAC	CTTCAGGA	ACTCAGACCG	1380
Qy	1429	TGCGTGA	TACTCAGGGGAGCCTGTGCCGTCGTATCCCTGTCC	CCAAAGGCAATCC	1488		
Db	1381	TGCGTGA	TACTCAGGGGAGCCTGTGCCGTCGTATCCCTGTCC	CCAAAGGCAATCC	1440		
Qy	1489	TTAAGA	AGTCTCGACAGCGTGAATCTGTATCTACTCTCT	CCAGAGCCACGCAAGTCTG	1548		
Db	1441	TTAAGA	AGTCTCGACAGCGTGAATCTGTATCTACTCTCT	CCAGAGCCACGCAAGTCTG	1500		
Qy	1549	GGGAA	ACTTTTGAAGGCCAGTGA	TGTGTTTGTGAAGTGGGAA	CCCGGTGAGCAAA	GTCTC	1608
Db	1501	GGGAA	ACTTTTGAAGGCCAGTGA	TGTGTTTGTGAAGTGGGAA	CCCGGTGAGCAAA	GTCTC	1568
Qy	1609	CACAGG	CTTTCAGGGGCTCTCTCC	CCACCGCAAGGGCATTCTG	CAAACTCAATG	GCAGATTCT	1668
Db	1561	CACAGG	CTTTCAGGGGCTCTCTCC	CCACCGCAAGGGCATTCTG	CAAACTCAATG	GCAGATTCT	1620
Qy	1669	CCGCGA	CAGCCTTAGAAGGCA	CTACCCCTTAGACA	CTTTTGGCTCCCTG	AGCCAACTG	1728
Db	1621	CCGCGA	CAGCCTTAGAAGGCA	CTACCCCTTAGACA	CTTTTGGCTCCCTG	AGCCAACTG	1680
Qy	1729	CTCTCC	CAATCTGCAAGCCCGGCGCCAGCCG	CCCTCAGAGGGGCTGTAGTGAAGAC	CAGCATCC	1788	
Db	1681	CTCTCC	CAATCTGCAAGCCCGGCGCCAGCCG	CCCTCAGAGGGGCTGTAGTGAAGAC	CAGCATCC	1740	
Qy	1789	TGTCCT	CGCAATCCCTTTGACCAATTTG	GAACCTTGCCGAAAC	CGTCTTCCCGAAAC	CCCACTGA	1848
Db	1741	TGTCCT	CGCAATCCCTTTGACCAATTTG	GAACCTTGCCGAAAC	CGTCTTCCCGAAAC	CCCACTGA	1800
Qy	1849	GGGCG	CTGTGTCTGTGGA	CAACCTGAGGGGGCTTTG	AGCAGCCTTCCCTCAG	AAAGTCTGA	1908
Db	1801	GGGCG	CTGTGTCTGTGGA	CAACCTGAGGGGGCTTTG	AGCAGCCTTCCCTCAG	AAAGTCTGA	1860
Qy	1909	AGCGAT	GGTGGCAAGAAATCTTTGGGGG	ATACCTGTTTTCTCTA	CAGACTGGCCAAAGG	1968	
Db	1861	AGCGAT	GGTGGCAAGAAATCTTTGGGGG	ATACCTGTTTTCTCTA	CAGACTGGCCAAAGG	1920	
Qy	1969	TGACTG	CAGCCTTAGACAAACCCCTAG	GAATCTGTCCAAAGCTC	AGCTGAGGAAAGGAGA	2028	
Db	1921	TGACTG	CAGCCTTAGACAAACCCCTAG	GAATCTGTGTGAGAGGATTC	2027		
Qy	2029	TGTGTG	CCCTAGTATGGGGT	TAGGCTCTGAGAGGGTTT	TGAGAGGAACC	2075	
Db	1981	TGTGTG	CCCTAGTATGGGGT	TAGGCTCTGAGAGGGTTT	TGAGAGGAACC	2027	

```

: Sequence 21 Application US/10332281
: Publication No. US20040126762A1
:
: GENERAL INFORMATION:
: APPLICANT: David W. Morris
: APPLICANT: Marc S. Malandro
: TITLE OF INVENTION: Novel Compositions and Methods in Cancer
: FILE REFERENCE: 529452001000
: CURRENT APPLICATION NUMBER: US/10/322,281
: CURRENT FILING DATE: 2002-12-17
: NUMBER OF SEQ ID NOS: 866
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 21
:
: LENGTH: 37278
:
: TYPE: DNA
:
: ORGANISM: Mus musculus
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(37278)
:
: OTHER INFORMATION: n = A,T,C or G
:
: US-10-322-281-21

```

Query March 1940, Conservative		66.8%; Score 19.8, 8;	DB 18;	Length 37278;
Best Local Similarity		99.9%;	Prod. No. 0;	
Matches 1940, Conservative		0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	955	CCGATGCTGTGACCTGATCCGGTGGCTGTTAATGGTGAACCCGACCGGTGGGCCACAC	1014	
Db	25129	CAGATGCTGTGGCTGTGATCCGGTGGCTGTTAATGGTGAACCCGACCGGTGGGCCACAC	25188	
Qy	1015	TGGAGATGTAGCCAGTCATTTGGTGGGTCACTGGGGTTTACACACCCGAGTGTGGGAAC	1074	
Db	25189	TGGAGATGTAGCCAGTCATTTGGTGGGTCAACTGGGGTTTACACACCCGAGTGTGGGAAC	25248	
Qy	1075	AGGAAGCCCGGTGAGGGGTGGGCAACCTAAGTGTGATCACTTTGGCCGGGCTCCATGGCCG	1134	
Db	25249	AGGAAGCCCGGTGAGGGGTGGGCAACCTAAGTGTGATCACTTTGGCCGGGCTCCATGGCCG	25306	
Qy	1135	ACTGTATACGTGCTCTCTCGCGCCCTCTCTGGAGATGGAGCCAGTGTGACCTTCT	1194	
Db	25309	ACTGTATACGTGCTCTCTCGCGCCCTCTCTGGAGATGGAGCCAGTGTGACCTTCT	25366	
Qy	1195	TCAAGCAGCAGCGTGGCCGGGAGGTGGAAGCATGTATCCTGGGCTGAGCGGCAATTCTC	1254	
Db	25369	TCAAGCAGCAGCGTGGCCGGGAGGTGGAAGCATGTATCCTGGGCTGAGCGGCAATTCTC	25428	
Qy	1255	TTAAGAAGTCCCGAAGGAGAAATGAATGGCTCAAAATGTGCAAGGTGACCGGGCTGAG	1314	
Db	25429	TTAAGAAGTCCCGAAGGAGAAATGAATGGCTCAAAATGTGCAAGGTGACCGGGCTGAG	25488	
Qy	1315	ATACCTCTTCTCGCCCTGGGCAAGCAGCTTAAAGCTTCCGAAGGCAATTCTCAAGAAA	1374	
Db	25489	ATACCTCTTCTCGCCCTGGGCAAGCAGCTTAAAGCTTCCGAAGGCAATTCTCAAGAAA	25548	
Qy	1375	AGTCTCTTACCTGTGCAAGGGAAGTACAGAGAGACCTTCAGGAATCAGACCGGTGCTG	1434	
Db	25549	AGTCTCTTACCTGTGCAAGGGAAGTACAGAGAGACCTTCAGGAATCAGACCGGTGCTG	25608	
Qy	1435	ATATCTCAAGGGGACGCTGTCTCTGCTGTATCCTTGCTCCCAAGAAAGGCAATCTTAA	1494	
Db	25609	ATATCTCAAGGGGACGCTGTCTCTGCTGTATCCTTGCTCCCAAGAAAGGCAATCTTAA	25668	
Qy	1495	AGTCTGACAGCGGTGAATCGTGTATCTATCACTCCCTCCAGAGCCAGCGAGTGTGGGAAC	1554	
Db	25669	AGTCTGACAGCGGTGAATCGTGTATCTATCACTCCCTCCAGAGCCAGCGAGTGTGGGAAC	25728	
Qy	1555	TCTTGAACGCGCAAGTATGTGTATTGTGAAGTGGGACCCCGTGAAGCAAGTCTTCCACAG	1614	
Db	25729	TCTTGAACGCGCAAGTATGTGTATTGTGAAGTGGGACCCCGTGAAGCAAGTCTTCCACAG	25788	
Qy	1615	CTTCAAGGCTCTCTCTTCCACCCGAAAGGCAATTCAAAATCMAATGGCAAGTTCTCCCGA	1674	
Db	25789	CTTCAAGGCTCTCTCTTCCACCCGAAAGGCAATTCAAAATCMAATGGCAAGTTCTCCCGA	25848	

QY 1675 CAGCTTAGAAGGCACTACCCCTTAGCACTTTGGCTCCCTGGACCACTGGCCCTCCCTCC 1734
DB 25849 CAGCTTAGAAGGCACTACCCCTTAGCACTTTGGCTCCCTGGACCACTGGCCCTCCCTCC 25908
QY 1735 ATCTGAGCCCGGCGCCAGCCGCTCAGGGGCTGTGAGTGAAGAGACAGATCCTGTCT 1794
DB 25909 ATCTGAGCCCGGCGCCAGCCGCTCAGGGGCTGTGAGTGAAGAGACAGATCCTGTCT 25968
QY 1795 CCGAGTCTTTGACCAATTGGAATTGGCTTGCCTGAACGTCTTCCGAAACCCCACTGAGGGGCT 1854
DB 25969 CCGAGTCTTTGACCAATTGGAATTGGCTTGCCTGAACGTCTTCCGAAACCCCACTGAGGGGCT 26028
QY 1855 GTGTGTCTGTGACAACCTGAGGGGGGCTTGAGCAACCTCCCTCGAAGGTCTGAAGCAT 1914
DB 26029 GTGTGTCTGTGACAACCTGAGGGGGGCTTGAGCAACCTCCCTCGAAGGTCTGAAGCAT 26088
QY 1915 GGTGGCAGGAATCCTTGGGGGATAGCTGCTTTCTCTGACAGACTGCAAGAGGTGACTG 1974
DB 26089 GGTGGCAGGAATCCTTGGGGGATAGCTGCTTTCTCTGACAGACTGCAAGAGGTGACTG 26148
QY 1975 CAGCTTAGAAGGCACTACCCCTTAGCACTTTGGCTCCCTGGACCACTGGCCCTCCCTCC 2034
DB 26149 CAGCTTAGAAGGCACTACCCCTTAGCACTTTGGCTCCCTGGACCACTGGCCCTCCCTCC 26208
QY 2035 CCTAGTATGGGGTATGGCTTGAAGGGGTTTGAGAGAACCTTGGGCTCGATTCCTCCAG 2094
DB 26209 CCTAGTATGGGGTATGGCTTGAAGGGGTTTGAGAGAACCTTGGGCTCGATTCCTCCAG 26268
QY 2095 TGAATAGAGTACATCAAGGGGCTCTACGTCTGACAGCTGACTGAACCTGAAGATGAGAGA 2154
DB 26269 TGAATAGAGTACATCAAGGGGCTCTACGTCTGACAGCTGACTGAACCTGAAGATGAGAGA 26228
QY 2155 AATCGCTTATGTGGAAGGAATGGAAACCTTGTGCTGCCGAGTGTATATGTGGGGTGG 2214
DB 26329 AATCGCTTATGTGGAAGGAATGGAAACCTTGTGCTGCCGAGTGTATATGTGGGGTGG 26388
QY 2215 CCTGAAGGTGCTACCTCTTTGTGCAATGATGATGCCATGACATTTCCCAACCCCTGT 2274
DB 26389 CCTGAAGGTGCTACCTCTTTGTGCAATGATGATGCCATGACATTTCCCAACCCCTGT 26448
QY 2275 CTCTGGCTGCACTTCAATAGTTTCTGTTCCATCAACCAAGGAGTTTGAACCTCTGA 2334
DB 26449 CTCTGGCTGCACTTCAATAGTTTCTGTTCCATCAACCAAGGAGTTTGAACCTCTGA 26508
QY 2335 CTCTGGCTGCACTTCAATAGTTTCTGTTCCATCAACCAAGGAGTTTGAACCTCTGTG 2394
DB 26509 CTCTGGCTGCACTTCAATAGTTTCTGTTCCATCAACCAAGGAGTTTGAACCTCTGTG 26568
QY 2395 CCACTCTGCTGCTGATCTCAAAAGACCTGGGGAAGCTCGAACCGCTGTTTGAACCTTCA 2454
DB 26569 CCACTCTGCTGCTGATCTCAAAAGACCTGGGGAAGCTCGAACCGCTGTTTGAACCTTCA 26628
QY 2455 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTATGATCTCAGAGACTTGAACCTTGA 2514
DB 26629 TCTCAAGGGGACAGATGCCCTGAGACCCCATCTTATGATCTCAGAGACTTGAACCTTGA 26688
QY 2515 GCTGTTCTTATGACCAAGATGTGATGATGCTGTTTCTCAGGCCAACGGGACCTAGA 2574
DB 26689 GCTGTTCTTATGACCAAGATGTGATGATGCTGTTTCTCAGGCCAACGGGACCTAGA 26748
QY 2575 ATGTGCTGACTTATTTATTTTGTGATCTCACTCTGTTTGTGTTTGTGTTTGTGTT 2634
DB 26749 ATGTGCTGACTTATTTATTTTGTGATCTCACTCTGTTTGTGTTTGTGTTTGTGTT 26808
QY 2635 TGTGTTGTTTGTGTTTATGATGATTTGTGCTTTCAATATGATGATGATGATGATG 2694
DB 26809 TGTGTTGTTTGTGTTTATGATGATTTGTGCTTTCAATATGATGATGATGATGATG 26868
QY 2695 GGGAACTCCACTGTGCCACTGAAGTTTATGTAAGAGAGATTTGGCAATGATGATGCT 2754
DB 26869 GGGAACTCCACTGTGCCACTGAAGTTTATGTAAGAGAGATTTGGCAATGATGATGCT 26928
QY 2755 CTATTCAGAGGGGGGTGGGGGGCTTTTCAATATGATGATGATGATGATGATGATG 2814

DB 26929 CTATTCAGAGGGGGGTGGGGGGCTTTTCAATATGATGATGATGATGATGATGATGATG 26988
QY 2815 TCTCAGTCCCTTACACCCCAAGGCTGGCCACCTCCCTCATCTTCACTGTGGCCAAA 2874
DB 26989 TCTCAGTCCCTTACACCCCAAGGCTGGCCACCTCCCTCATCTTCACTGTGGCCAAA 27048
QY 2875 AAAAAAAAAAAAAAAAAAAAAA 2896
DB 27049 AAAAAAAAAAAAAAAAAAAAAA 27070
RESULT 6
US-10-343-514-27
Sequence 27, Application US/10343514
Publication No. US20040132025A1
GENERAL INFORMATION:
APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEBEYRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 2026
TYPE: DNA
ORGANISM: RAT
US-10-343-514-27
Query Match 61.0%; Score 1771.4; DB 18; Length 2026;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 136; Indels 5; Gaps 3;
QY 49 GTGACCTTGAAGCCCGGCTCAGCGCGCTGCTACTGCTGCGGACCCACTTCAACTTC 108
DB 1 GTGACCTTGAAGCCCGGCTCAGCGCGCTGCTACTGCTGCGGACCCACTTCAACTTC 60
QY 109 GCGGTCCCGCACATGAGATGCTGTGCTTACTTCAAGCGCCGAGCCAGGCTCCCTCG 168
DB 61 GCGGTCCCGCACATGAGATGCTGTGCTTACTTCAAGCGCGCCGAGGAACTTGTCCCTCG 120
QY 169 CCTCCGCGCTGGCTGGAGAGCGCGCGCGCTGGCGGAGCGGGCTCATCAAGTCCGCTA 228
DB 121 CTTCCGCGCTGGCACAGAGAGCGCGCGCGCTGGCGGAGCGGGCTCATCAAGTCCGCTA 180
QY 229 AACCTTGAAGAAAGCAAGCGCTGAGAGCGGACCATCAACAACAACCTTGGCGACC 288
DB 181 AACCTTGAAGAAAGCAAGCGCTGAGAGCGGAGCGGACCATCAACAACAACCTTGGCGACC 240
QY 289 GCTACGATGCTCTGAGAGCGCTGGCAAGGCACTTACGGGAAGTGAAGAGCAAGAG 348
DB 241 GCTACGATGCTCTGAGAGCGCTGGCAAGGCACTTACGGGAAGTGAAGAGCAAGAG 300
QY 349 AGAGCTCGGGGGCTGGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 408
DB 301 AGAGCTCGGGGGCTGGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAGATGAGC 360
QY 409 AGGATCTGTGCAATACGAGAGGAGATTTGAGATCAATGCTTCACTCAACCAACCCCA 468
DB 361 AGGATCTGTGCAATACGAGAGGAGATTTGAGATCAATGCTTCACTCAACCAACCCCA 420
QY 469 TCATTCATCATCAAGAGTGTGGAATATGACAGCAAGATTTGATGATGATGATGATG 528

Db 421 TCATTGCAATCCATGAAAGTGTGGAAGACAGCAAGCAAGATTGTATGTCATGGAATACG 480
Qy 529 CCAGCCGAGGAGATCTGTATGATTAATCATCATGAGGCGGCAAGCTGATGAGGAGGAG 588
Db 481 CCAGCCGAGGAGATCTGTATGATTAATCATCATGAGGCGGCAAGCTGATGAGGAGGAG 540
Qy 589 CCAGGCAATTTCTTCCGACAGATCGTCTGCGCTGCACTATGTCACACAGAGCGGATCG 648
Db 541 CCAGGCAATTTCTTCCGACAGATCGTCTGCGCTGCACTATGTCACACAGAGCGGATCG 600
Qy 649 TTCAACCAAGATCTCACTGAGGAAAAATCTCTTCAATGATGCAATGAAAAATCAAGATTG 708
Db 601 TTCAACCAAGATCTCACTGAGGAAAAATCTCTTCAATGATGCAATGAAAAATCAAGATTG 660
Qy 709 CTGACTTTGGGCTCTCCAACTGATACCAAGGCAAGGCTCTGACAGGCTTCTGTGAGGA 768
Db 661 CTGACTTTGGGCTCTCCAACTGATACCAAGGCAAGGCTCTGACAGGCTTCTGTGAGGA 720
Qy 769 GGCCTCTCTACGCTCTGCTGAGATGATCAACGAGAAAGCCCTATGAGGCGCCAGAGGTG 828
Db 721 GGCCTCTCTACGCTCTGCTGAGATGATCAACGAGAAAGCCCTATGAGGCGCCAGAGGTG 780
Qy 829 ACAGCTGCTCTGAGGCGGCTCTCTGATCATCTGTGATGATGCAACATGCGCTTTGACG 888
Db 781 ACAGCTGCTCTGAGGCGGCTCTCTGATCATCTGTGATGATGCAACATGCGCTTTGACG 840
Qy 889 GGAAGATCAATPAAACCTGATGAGCAATCAATACGAGGAGCTTACCTGAGGCGGCCA 948
Db 841 GGAAGATCAATPAAACCTGATGAGCAATCAATACGAGGAGCTTACCTGAGGCGGCCA 900
Qy 949 AGCGCTCGATGCTGTGAGGCTGATCGGAGGCTGTTAATGATGAAACCCGCGGAGG 1008
Db 901 AGCGCTCGATGCTGTGAGGCTGATCGGAGGCTGTTAATGATGAAACCCGCGGAGG 960
Qy 1009 CCAACTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1068
Db 961 CCAACTGAGAGATGTAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 1069 GGAAGACGAAAGCCCTGCTGAGAGGATGAGGAGCCCTTATGATGATGATGATGATG 1128
Db 1021 GGAAGACGAAAGCCCTGCTGAGAGGATGAGGAGCCCTTATGATGATGATGATGATG 1080
Qy 1129 TGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1188
Db 1081 TGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1189 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGAGACACTGTATCTGTGAGGCTGAGCGGAC 1248
Db 1141 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGAGACACTGTATCTGTGAGGCTGAGCGGAC 1200
Qy 1249 ATTCTCTTAAGAGATCCCGAAAGAAATGATCATGAGCTCAAAATCTGCAAGATGATCC 1308
Db 1201 ATTCTCTTAAGAGATCCCGAAAGAAATGATCATGAGCTCAAAATCTGCAAGATGATCC 1260
Qy 1309 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
Db 1261 TTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy 1369 AGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1428
Db 1321 AGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Qy 1429 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1488
Db 1381 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Qy 1489 TTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1548
Db 1441 TTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 1549 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1608

Db 1501 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Qy 1609 CAGAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1668
Db 1561 CAGAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1617
Qy 1669 CCGGACAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1728
Db 1618 CCGGACAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1677
Qy 1729 CCGGACAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1788
Db 1678 CCGGACAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1737
Qy 1789 TGCTCTCCAGGCTCTTTCAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1848
Db 1738 TGCTCTCCAGGCTCTTTCAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1797
Qy 1849 GGGGCTGT 1908
Db 1798 GGGGCTGT 1857
Qy 1909 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1968
Db 1858 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1917
Qy 1969 TGACTGAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 2028
Db 1918 TGACTGAGGCTTACGAGGCTCTCTCTCCACGCAAGGAGCAATTCATCAATGAGCAAGTTCT 1977
Qy 2029 TGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2088
Db 1978 TGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2026

RESULT 7
US-09-963-159-1
Sequence 1, Application US/09963159
Patent No. US2002007712A1
GENERAL INFORMATION:
APPLICANT: CURTIS, RORY A.J.
APPLICANT: GALVIN, KATHERINE M.
TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
FILE REFERENCE: 10147-5001
CURRENT APPLICATION NUMBER: US/09/963,159
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/234,922
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3353
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (3268)
NAME/KEY: unsure
LOCATION: (3270)
NAME/KEY: unsure
LOCATION: (3272)
US-09-963-159-1

Query Match 47.5%; Score 1379; DB 9; Length 3353;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;
Qy 82 CTACTGTGCGCCGACCACTTCACCTGCGGATCCCGCAGCAATGAGATCGATGCTTAC 141
Db 116 CTACTGTATTCCTCGCGCCGCTTGTCTACCTCTGCTGCGCATGAGATTCCTGTTTCG 175
Qy 142 TCAGCGCCGAGCGAGGCTTCCTGCGCTTCGCGCTTCGAGAGAGCGCCGCGCGC 201

Db 176 CGGCGCGCTCCGGCCCCCACTCCCTCGGCGGAGAGCTA-----GCCCGGGCGC 223
Qy 202 TGGCGGACGGGCTCATTAAGTCGGCTAACTCTGATGAGAGACAGCGGTGAAGCGC 261
Db 224 TGGGGGAGGGGCTGATCAAGTCGGCCAAAGCCCTTAAGAGAACAGCGGTGAAGCGC 283
Qy 262 ACCATCAAAACAAACCTGCGGACCGCTACGAGTTCTTGAGAGACGCTGGGCAAGGCA 321
Db 284 ACCACCAAGGACAACTTGGGCAACCGCTACGAGTTCTTGAGAGACCTGGGCAAGGCA 343
Qy 322 CTTACGGGAAAGTGAAGAGACGAGAGAGCTCGGGCGCTCTGTGGCCATCAAGTCA 381
Db 344 CTTACGGGAAAGTGAAGAGAGCGGAGAGCTCGGGCGCTCTGTGGCCATCAAGTCA 403
Qy 382 TCAGAGAAAGCAAAATCAAGATAGAGAGATCTGTGCATACACGAGGAGATTAGA 441
Db 404 TCGGAGAGGACAAATCAAGATAGAGAGATCTGATGCAATCGAGGAGAGATTAGA 463
Qy 442 TCAGTCTTCACTCAACACCGCCCACTCATTCATTCATGAGTGTGAGAAATGCA 501
Db 464 TCATGTATCATCTCAACACCGCTCATCATTCATTCATGAGTGTGAGAAACGCA 523
Qy 502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 524 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
Qy 562 AGCGGCAACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
Db 584 AGCGGCAACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 643
Qy 622 TGCACTACTGCAACAGAGCGGATCGTTACCGAGATCTCAAGCTGAGAAACCTCT 681
Db 644 TGCACTACTGCAACAGAGCGGATCGTTACCGAGATCTCAAGCTGAGAAACCTCT 703
Qy 682 TAGATGCAATGAGAAATCAAGATGATGATGATGATGATGATGATGATGATGATG 741
Db 704 TAGATGCAATGAGAAATCAAGATGATGATGATGATGATGATGATGATGATGATG 763
Qy 742 GCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTCTACGCTCGCTGAGATGATGATG 801
Db 764 GCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTCTACGCTCGCTGAGATGATGATG 823
Qy 802 GGAAGCCCTAATGAGGCGCAGAGGTGACAGCTGCTCTGTGGGCTCTCTGATCAATC 861
Db 824 GGAAGCCCTAATGAGGCGCAGAGGTGACAGCTGCTCTGTGGGCTCTCTGATCAATC 883
Qy 862 TGGTGCATGAGCAACATGCTTTGACCGGACAGATCAATAAACATGCTGAGAGCAAT 921
Db 884 TGGTGCATGAGCAACATGCTTTGATGGGATGACATTAAGATCTGATGAGAAACAT 943
Qy 922 GTAACGGGCTTACCTGAGCGCGCCCAAGCGCTCGATGCTGTGCTGATCCGATG 981
Db 944 GTAACGGGCTTACCTGAGCGCGCCCAATTAACCTCTGATGCTGTGCTGATCCGATG 1003
Qy 982 TGTTAATGAGAACCCCAACCGCTCGGACCACTGAGAGATGATGAGAGATGATGAG 1041
Db 1004 TGTTAATGAGAACCCCAACCGCGGACCACTGAGAGATGATGAGAGATGATGAG 1063
Qy 1042 TCAACTGGGCTTACCAACCGGAGTCCGGGAAACAGAAAGCCCTGAGAGGATGAG 1101
Db 1064 TCAACTGGGCTTACCAACCGGAGTCCGGGAAACAGAGAGCTCCGATGAGAGATG 1123
Qy 1102 CTAGTGTGATCTTTGCGCGGCTCTCATGCGGAGCTGTGATACGTGCTCTCGCGCC 1161
Db 1124 CTAGTGTGATCTTTGCGCGGCTCTCATGCGGAGCTGTGATACGTGCTCTCGCGCC 1183
Qy 1162 TCTGAGAGATGAGAACAGTGTGAGCTTTCTTCAAGCAGACAGTGCCTGGAGAGTGA 1221
Db 1184 TCTGAGAGATGAGAACAGTGTGAGCTTTCTTCAAGCAGACAGCTGTGTGGAGAA 1243
Qy 1222 GCACTGATCTGGGCTGAGAGCGGACATTTCTTAAAGAGTCCGAAAGAGATGACA 1281
Db 1244 GCACTGATCTGGGCTGAGAGCGGACATTTCTTAAAGAGTCCGCAAGAGATGACA 1303

Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGAGATACCTTCTCGCCCTGGCAAGCA 1341
Db 1304 TGGCCAGCTCTCCACAGTGAACAGGCTGATGACATGCTCCATCGCTGGCAAGCA 1363
Qy 1342 GCCCTTAAGCTTCCGAAAGGCACTTCAAGAAAGTCTCTCACTCGTCAAGGAGGATC 1401
Db 1364 ACCTCAAGCTGCCAAGAGGCTTCTCAAGAAAGTGTGATGATCTGTGCAAGAGGATC 1423
Qy 1402 AGGAGACCTCTCAGAACTCAGACCGATGCTGATATCTCAGAGGCAAGCTGTCTG 1461
Db 1424 AGGAGACCTCTCAGAGCTCAGCCCAATCTCTGAGCCAGGCAAGGCTGTCTG 1477
Qy 1462 TATCTGCTCTCCAGAAAGGCACTCTTAAGAGTCTGACAGCTGATCTGTGTTACT 1521
Db 1478 ---CCCTGCTCTCCAGAAAGGCACTCTCAAGAAAGCCCAAGCGGATCTGTGTTACT 1534
Qy 1522 ACTCTCTCCAGAGCCAGGATCTGAGGAACTCTTACAGCGCAGTATGATGTTGTA 1581
Db 1535 ACTCTCTCCAGAGCCAGGATCTGAGGAACTCTTACAGCGGATGATGTTGTA 1594
Qy 1582 GTGGGACCCCTGAGAGCAAGATCTCAAGGCTTCAAGGCTCTCTTCAAGCGAGG 1641
Db 1595 GTGGGATCTCCAGAGCAAGATCTCAAGGCTTCAAGGCTCTCTTCAAGCGAGG 1654
Qy 1642 GCATTTCAAACTCAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1701
Db 1655 GCATTTCAAACTCAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1714
Qy 1702 CTTTGGCTCTCCAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1761
Db 1715 CTTTGGCTCTCCAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1774
Qy 1762 CAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
Db 1775 CAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1834
Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 1881
Db 1835 CTGAAGCTCTTCCGAAACCCCACTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 1894
Qy 1882 TTGAGCAAGCTCTCTCAGAAAG-----GTCTGAAGGATGATGAGCAAGATCT 1929
Db 1895 TTGAGCAAGCTCTCTCAGAAAGGCTCTGAGAGCTGATGAGGCTGTGAGAGATCT 1954
Qy 1930 TGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1989
Db 1955 TGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2014
Qy 1990 CCTTGAAGATCTGCTCAAGCTCAAGCTGAGAGAGGATGATGATGATGATGATGATG 2049
Db 2015 CACTGAGGCTCTCTCAAGCTCAAGCTGAGAGAGGATGATGATGATGATGATGATG 2073
Qy 2050 GCTCTGAGAGGCTTTCAGAGGAACTCTGATGATGATGATGATGATGATGATGAT 2092
Db 2074 GCTCTGAGAGGCTTTCAGAGGAACTCTGATGATGATGATGATGATGATGATGAT 2116

RESULT 8
US-10-423-543-43
; Sequence 43, Application US/10423543
; Publication No. US2004058355A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Lieberman, Rosana K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Myoung

APPLICANT: Williamson, Mark J.
APPLICANT: Siles-Santiago, Inmaculada
APPLICANT: Bandaru, Rajasekhara
TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 36555 OR 593 MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: MP103-0230NMIM
CURRENT APPLICATION NUMBER: US/10/423, 543
CURRENT FILING DATE: 2003-04-25
PRIORITY APPLICATION NUMBER: US 10/278, 036
PRIORITY FILING DATE: 2002-10-22
PRIORITY APPLICATION NUMBER: US 09/711, 216
PRIORITY FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: US 60/205, 447
PRIORITY FILING DATE: 2000-05-19
PRIORITY APPLICATION NUMBER: US 10/012, 055
PRIORITY FILING DATE: 2001-11-13
PRIORITY APPLICATION NUMBER: US 60/248, 325
PRIORITY FILING DATE: 2000-11-14
PRIORITY APPLICATION NUMBER: US 10/003, 690
PRIORITY FILING DATE: 2001-11-15
PRIORITY APPLICATION NUMBER: US 60/248, 893
PRIORITY FILING DATE: 2000-11-15
PRIORITY APPLICATION NUMBER: US 09/797, 039
PRIORITY FILING DATE: 2001-02-28
PRIORITY APPLICATION NUMBER: US 60/186, 061
PRIORITY FILING DATE: 2000-02-29
PRIORITY APPLICATION NUMBER: US 10/217, 168
PRIORITY FILING DATE: 2002-08-12
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 119
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 3353
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(3353)
OTHER INFORMATION: n = A,T,C or G
US-10-423-543-43

Query Match 47.5%; Score 1379; DB 17; Length 3353;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;

QY 82 CTACTGTCGCCGACCACTTCACTCGCGGTCCCGCACCATGAGTGGCTTAC 141
DB 116 CTACTGATTCCTCCGCGCCCTTCTCACTCGCTGCGCATGAGTGGCTTAC 175
QY 142 TCCAGCGCCGAGCGCAAGTCTCTGCGCTCCGCTCGGCTCGAGAGCGCCGCGC 201
DB 176 CGCGCGCTCTCGCGCCCACTCTCTCGCGCGAGAGTA-----GCCGCGCGC 223
QY 202 TGGGGAAGCGGCTCATGAGTGGCTTAACCTCTGAGTGAAGAAGCGGCTGAAGGCG 261
DB 224 TGGGGAAGGCTGATCAAGTGGCTCAAGCCCTTAATGAAGACAGCGGTGAAGGCG 283
QY 262 ACCATCAAAACAACTCTCGGCGCAAGCTTCAAGTCTTGAAGCGCTGGGCAAGGCA 321
DB 284 ACCACCAAGGCACTCTGCGGCAAGCTTCAAGTCTTGAAGCGCTGGGCAAGGCA 343
QY 322 CTAAGGGAAGTGAAGAGGCAAGAGAGTCTGGGCGTCTGGTGGCATCAAGTCA 381
DB 344 CTAAGGGAAGTGAAGAGGCGGGAAGAGTCTGGGCGTCTGGTGGCATCAAGTCA 403
QY 382 TCAGGAAGACCAAAATCAAGATGAGAGGATCTGTGCAATCAAGGGAAGTGAAG 441
DB 404 TCAGGAAGACCAAAATCAAGATGAGAGGATCTGTGCAATCAAGGGAAGTGAAG 463
QY 442 TCATGTCTTCACTCAACCAAGCCCAATCTGCGATTCATGAAGTGTGAAGATGCA 501

DB 464 TCATGTCTTCACTCAACCAAGCCCAATCTGCGATTCATGAAGTGTGAAGATGCA 523
QY 502 GCAAGATTTGATTTGATGAGAGTATGACAGCCGAGCGATCTGATGATTTACATG 561
DB 524 GCAAGATTTGATTTGATGAGAGTATGACAGCCGAGCGATCTTATGATTTACATG 583
QY 562 AGCGGCAAGGCTGAGTGAAGCGGAGCGGCAAGTCTTCTTCCAGATGCTGTGCGC 621
DB 584 AGCGGCAAGGCTGAGTGAAGCGGAGCGGCAAGTCTTCTTCCAGATGCTGTGCGC 643
QY 622 TGCATCTATCCGACCAAGAGGATCGTTACAGGAGTCTCAAGCTGGAAGAAATCCCTC 681
DB 644 TGCATCTATCCGACCAAGAGTCTTCAAGGAGTCTCAAGCTGGAAGAAATCCCTC 703
QY 682 TAGATGCAATGGAAGATCAAGATTTGATGAGTCTTCTTCCAGATGCTGTGCGC 741
DB 704 TAGATGCAATGGAAGATCAAGATTTGATGAGTCTTCTTCCAGATGCTGTGCGC 763
QY 742 GCAAGTCTTCCAGAGCTTCTGAGGAGCCCTCTTCAAGCTGCTGAGATGATCAAG 801
DB 764 GCAAGTCTTCCAGAGCTTCTGAGGAGCCCTCTTCAAGCTGCTGAGATGATGATG 823
QY 802 GGAAGCCCTATGAGGAGCCGAGAGTGAAGAGTCTTCTGAGGCTTCTCTGATCAATCC 861
DB 824 GGAAGCCCTATGAGGAGCCGAGAGTGAAGAGTCTTCTGAGGCTTCTCTGATCAATCC 883
QY 862 TGGGCAATGGAAGATCAAGATTTGATGAGTCTTCTTCAAGATGATGATGATGATCA 921
DB 884 TGGGCAATGGAAGATCAAGATTTGATGAGTCTTCTTCAAGATGATGATGATGATGATCA 943
QY 922 GTAAAGGAGGCTTACCGTGAAGCGGCGCAAGCGTTCAGTCTGAGCTGATCCGATGAGC 981
DB 944 GCAAGGAGGCTTACCGTGAAGCGGCGCAAGCGTTCAGTCTGAGCTGATCCGATGAGC 1003
QY 982 TGTTAATGTAAGATCCCAAGGCTGAGGCTGAGGATGAGATGAGATGATGATGATGAG 1041
DB 1004 TGTTAATGTAAGATCCCAAGGCTGAGGCTGAGGATGAGATGAGATGAGATGAGATGAG 1063
QY 1042 TCAACTGAGGATTCACCAAGGAGTGGGGAAGAGAGGCTGAGGAGGAGGAGGAGGAGC 1101
DB 1064 TCAACTGAGGATTCACCAAGGAGTGGGGAAGAGAGGCTGAGGAGGAGGAGGAGGAGC 1123
QY 1102 CTAAGTGTGATTTTGGCGGCGCTTCAAGTGGGAGTCTGATGATGATGATGATGATGATG 1161
DB 1124 CTAAGTGTGATTTTGGCGGCGCTTCAAGTGGGAGTCTGATGATGATGATGATGATGATG 1183
QY 1162 TCTTGAAGATGAGAGGAGTGTGAGCTTCTTCAAGAGAGAGTGTGAGAGTGTGAG 1221
DB 1184 TCTTGAAGATGAGAGGAGTGTGAGCTTCTTCAAGAGAGAGTGTGAGAGTGTGAGAGTGTGAG 1243
QY 1222 GCACTGTATCTGAGGCTGAGAGGAGTGTGAGCTTCTTCAAGAGAGAGTGTGAGAGTGTGAG 1281
DB 1244 GCACTGTATCTGAGGCTGAGAGGAGTGTGAGCTTCTTCAAGAGAGAGTGTGAGAGTGTGAG 1303
QY 1282 TGGGCAAAATCTGAGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1341
DB 1304 TGGGCAAAATCTGAGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1363
QY 1342 GCTTAAAGCTTCCGAAAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1401
DB 1364 ACCTCAAGTGTGAGAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1423
QY 1402 AGAGAGAGCTTCAAGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1461
DB 1424 AGAGAGAGCTTCAAGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1477
QY 1462 TATCTCTGCTCCGAAAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1521
DB 1478 TATCTCTGCTCCGAAAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1543
QY 1522 ACTCTCTGCTCCGAAAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1581
DB 1535 ACTCTCTGCTCCGAAAGGAGTGTGAGAGGAGTGTGAGCTTCTTCAAGAGAGTGTGAGAGTGTGAG 1594

OY	1582	GTGGGGACCCCGTAGGACGAAAGTCTCCACAGGCTTACAGGGCTCTCCCTCCACCGGCAAG	1641
Db	1595	GTGGGGATCCCAAGAGACGAAAGCTCCGGCAGCTTCAAGGCTCTCTCTTCATCGCAAG	1655
OY	1642	GCATTCTCAAACTCAATGCGAAATTTCTCCGACAGCCTTAGAAGGCACTAACCCCTAGCA	1701
Db	1655	GCATCTCTCAAACTCAATGCGAAATTTCTCCGACAGCCTTAGAAGCTCCCGGCCCCCA	1714
OY	1702	CCTTTGGCTCCCTGGACCAACTGGGCTCCTCCCACTCTGACAGCCGGGCCAGCCGCCCT	1761
Db	1715	CCTTGGCTCCCTGGATGAACCTCGCCCCCACTCGCCCCCTGGCCCGGCGCAGCTGACCTT	1774
OY	1762	CAGGGGCTGTGAGTAGGAGACAGCATCTGTCTCTCCGAGTCTCTTTGACCAATTGACTTGC	1821
Db	1775	CAGGGGCTGTGAGCGAGGACAGCACTCTGTCTCTGAGTCTCTTTACAGCTGAGCTTGC	1834
OY	1822	CTGAACGTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGTGACCAACTGAGGGGCG	1881
Db	1835	CTGAACGGCTCTCCCAAGAGCCCCCACTGCGGGGCTGTGTGTGTGTGACCAACTCAGGGGCG	1894
OY	1882	TTGAGCAGCCTCCCTCAGAG-----GTCTGAAGGATGTGTGGCAGGAAATCTT	1929
Db	1895	TTGAGAGACCCCTCCCAAGAGGCCCTGTGAAGCTGTCTGAAGGGCTGGCGGAGGATCTT	1954
OY	1930	TGGGGGATAGTGTCTTTTCTCTGACAGACTGCGCAAGAGGTGACTGACGCTTACAGACAAG	1989
Db	1955	TGGGGGACAGGCTGCTTTTCCCTGACAGACTGCGCAGAGGTGACAGCACTTACCGACAGG	2014
OY	1990	CCCTTAGAATCTGCTCAAAAGCTCAGCTGAGGAGGAGATGGTGCCTTAGTATGGGGTAG	2049
Db	2015	CACTAGAGGGTCTGCTCAAAAGCTCACCTAAGTGAAGTAGGCAATTGGCCCCAG--CCGGGTAG	2073
OY	2050	GCTCTGAGAGGGTTTGCAGAGGAACCTCGGGTCCGATTCTCC 2092	
Db	2074	GCTCTCAGATGCACTGCTGGTTGACCCCGAGGGGAGATGCTTTC 2116	
RESULT 9			
US-10-322-281-25			
Sequence 25, Application US/10322281			
Publication NO. US20040126762A1			
GENERAL INFORMATION:			
APPLICANT: David W. Morris			
APPLICANT: Marc S. Malandro			
TITLE OF INVENTION: Novel Compositions and Methods in Cancer			
FILE REFERENCE: 529452001000			
CURRENT APPLICATION NUMBER: US/10/322,281			
CURRENT FILING DATE: 2002-12-17			
NUMBER OF SEQ ID NOS: 866			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 25			
LENGTH: 3404			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-10-322-281-25			

Query Match	47.4%	Score 1375.8	DB 18	Length 3404
Best Local Similarity	81.7%	Pred. No. 0		
Matches 1652	Conservative	0	Mismatches 337	Indels 34
			Gaps	4
QY	CTACGCGCCGACCCACCTCACTCGCGGATCCCCGACGATGAGTGGGTGGCTTAC	141		
Db	87 CTATCTGATYCCCCCTGCGCCCTTGCTCACTCTGCTGCGATGAGTGGGTGGTTCG	146		
QY	142 TCCAGCGCCGAGCGCAGGCTCCTCTGCGGCTCGGCGCTGGCTCGAGAGCGCCGCGCCG	201		
Db	147 CGCGGCGCTCGGCGCCCACTCTCTCGGCGGAGACTY-----GCCGCGCCG	194		
QY	202 TGCGGAGCGGCTCATCAAGTGGCTTAACCTTGATGAAGAGCAGCGGTGAAGCGC	261		
Db	195 TGGGGAAGGGCTGATCAAGTGGCCCAAGCCCTTATTAAGAAGCAGCGGTGAAGCGC	254		

OY	262	ACCATCAAAACAAACCTGCGGGACCGCTACAGAGTTCTTGAGACGCTGGGCAAGGGCA	321
Db	255	ACCAACACAAACAACCTGCGGACCGCTACAGAGTTCTTGAGACCTGGGCAAAAGCA	314
OY	322	CTTACGGGAAGTGAAGAGGCAAGAGAGCTCGGGGCTCTGGTGGCCATCAAGTCA	381
Db	315	CTTACGGGAAGTGAAGAGAGCGCGGGAGAGCTCGGGGCGCTGGTGGCCATCAAGTCA	374
OY	382	TCAGAAAGACAAATTCAAAGATGACAGAGATCTGCTGCACATACGAGGGAGATTGGA	441
Db	375	TCAGAAAGACAAATTCAAAGATGACAGAGATCTGATGCAATACGAGGGAGATTGGA	434
OY	442	TCATGTCCTTCAACCAACCCCAACATCTTGCCATCATGAAAGTGTGGAATAGCA	501
Db	435	TCATGTCATCTCAACCAACCCCTACATCATTTGCCATCATGAAGTGTGGAAGACGA	494
OY	502	GCAAGATTGTATGTCAATGAGATATGCCAGCCGAGCGGATCTGTATGATTAATCAAGTG	561
Db	495	GCAAGATTGTATGTCAATGAGATATGCCAGCCGAGCGGAGACTTTATACATACACG	554
OY	562	AGCGGCCACGGGCTGATGAGCGGGACCGCAGGCAATTTCTTCCGACAGATCGTGTGCGC	621
Db	555	AGCGGCAGAGCTCAGTACGCGGCCAAGCTAAGGCAATTTCTTCCGGCAGATCGTGTGCGC	614
OY	622	TGCACCTACTGCAACAGAGCGGGATTCGTCAACGAGATCTCAACCTGGAAAAATCTCTC	681
Db	615	TGCACATATTCGATCAACAGAGGATTTGTCCACCGAGATCTCAACCTGGAAAAATCTCTC	674
OY	682	TGATATGCCAATGAGAAACATCAAGATTGCTGACTTTGGGCTCTCCAACTGTACACAAG	741
Db	675	TGATATGCCAATGGGAAATCAAGATTGCTGACTTGGGCTCTCCAACTGTACACATCAG	734
OY	742	GCAAGTTCCTCAGACGTTCTGTGGGAGACCTCTCTACGCTGTGACTGATAGTCAACG	801
Db	735	GCAAGTTCCTGCAACATCTGTGGGAGACCCCTCTAATGCTGTCCACAGAGATTGTCAATG	794
OY	802	GGAAGCCCTAATGTGGGCGCCAGAGGTGACAGCTGTCTGTGGGGTCTCTCTGTACATCC	861
Db	795	GGAAGCCCTTACACAGGCCCAAGAGGTGACAGCTGTCTCTGGGGTCTCTCTTACATACC	854
OY	862	TGATGCAATGAGCAACATGCTTTTACCGGGACGAGATCATTAACACTGGTGAAGCAAAATCA	921
Db	855	TGATGCAATGAGCAACATGCTTTTATGAGGACATGACCATTAAGATCTTATGTAACACATCA	914
OY	922	GTAACGGGGCTTACCGTGAAGCCGCCAAGCCGTCCGATGCTCTGTGGCTGTATCCGGTGGC	981
Db	915	GCAACGGGGCTTACCGGAGAGCCATCTTAACCTCTGATGCTGTGGCTGTATCCGGTGGC	974
OY	982	TGTTAATGTGTAAACCCCAACCGGTGGGCGACATGTGAGAGATGACCATGATGTGTGGG	1041
Db	975	TGTTAATGTGTAAACCCCAACCGGCGGCGCACCTGTGAGAGATGTGGCCATGATGTGTGGG	1034
OY	1042	TCAACTGGGGGTTACACCAACCGGAGATCGGGGGAACAGAAAGCCCTGCGTAGAGGTGGCAC	1101
Db	1035	TCAACTGGGGGTTACGCAACCGAGATGGGAGAGCAGAGGCTCCGCAATAGGGGTGGCAC	1094
OY	1102	CTAATGTGTGACTTTGGCGGGGCTCCATATGGCGGACTGTGTTAATCTGTGCTCTCGCGCCCC	1161
Db	1095	CTGCGCATGACTCTGCCCGCGGCTCCATATGCTGTACTGGCTCCGGGCTTCTCTCCGCCCC	1154
OY	1162	TCCTGAGAAATGAGACCAAGGTGTGCAGCTTTCAAGCAGACGTCGCCGGGAGGTGGAA	1221
Db	1155	TCCTGAGAAATGAGGGGCAAGGTGTGACAGCTTTCTTCAAGCAGACGCTGTGGGGGAA	1214
OY	1222	GCACTGTAACTTGGGCTGGAGCGGGAATTTCTCTTAAGAAATCCCGAAAGGAATGACA	1281
Db	1215	GCAACACACCTTGGGCTTGAAGCGGCAAGATTCGCTCAAGAAATGCCGAGAGGATGACA	1274
OY	1282	TGGCTCAAAATCTGTCAAGTGAACCGGCTGAGATACCTCTTCCGCGCTGGCAAGACA	1341
Db	1275	TGGCCCAAGCTCTTCCACAGTGAACGGCTGTATGACATGCCCATGCGCTTGGCAAGACA	1334
OY	1342	GCCCTTAAGCTTCGAAAGGCAATCTCAAGAAAAATGCTCTTACCTGTGACGGGAGGTAC	1401

```

Db 1335 ACCCTGAGCTGCCAAGGAGGATTTCTCAAGAGAGAGGTGTGAGCTCTGCAAGAGGGGTAC 1394
QY 1402 AGGAGAGACCTCTAGAGACTGAGACCGGAGCTGTATCTCCAGGGGAGCTGTCTCTCTG 1461
Db 1395 AGAGAGACCTCTGAGAGCTGAGCCCAATCTCTGAGAGCCAGGGAGAGCTGCTCTCTG 1448
QY 1462 TATCTCTGCTCCAGAGAGAGGCACTCTTAAGAGTGTGACAGCTGAACTTGTGTACT 1521
Db 1449 ---CCCTGCTCTCCAGAGAGGAGTCTCAAGAGAGCCCAAGAGAGAGTGTGTGTACT 1505
QY 1522 ACTCTCTCTCCAGAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1581
Db 1506 ACTCTCTCTCCAGAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1565
QY 1582 GTGGGAGACCCGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1641
Db 1566 GTGGGAGATCCAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1625
QY 1642 GCATCTCAAACTCAATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1701
Db 1626 GCATCTCAAACTCAATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
QY 1702 CTTTGTGCTCTGAGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1761
Db 1686 CTTTGTGCTCTGAGATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1745
QY 1762 CAGGGGCTGTGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1821
Db 1746 CAGGGGCTGTGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1805
QY 1822 CTGAACGTCTTCCGAGAACCCCACTGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1881
Db 1806 CTGAACGGCTCTCCAGAGAGCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1865
QY 1882 TTGAGAGACCTCTCTCAAGAG-----GTCTGAAGCAGTGTGTGTGTGTGTGTGTGT 1929
Db 1866 TTGAGAGAGCCCTCTCAAGAGAGCCCTGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1925
QY 1930 TGGGGGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1989
Db 1926 TGGGGGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1985
QY 1990 CCTTGAAGATCTGTCTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2049
Db 1986 CACTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2044
QY 2050 GCTCTGAGAGGTTTTCAGAGAGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2092
Db 2045 GCTCTGAGATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2087

RESULT 10
US-10-370-715B-639
; Sequence 639, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
;   Patin Docket Preview
;   APPLICANT: BODARY, SARAH C.
;   APPLICANT: CLARK, HILARY
;   APPLICANT: BRISDELL, HUNTE
;   APPLICANT: JACKMAN, JANET
;   APPLICANT: SCHOENFELD, JILL R.
;   APPLICANT: WILLIAMS, P. MICKEY
;   APPLICANT: WOOD, WILLIAM I.
;   APPLICANT: WU, THOMAS D.
;   TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
;   TITLE OF INVENTION: Related Diseases
;   FILE REFERENCE: P19481-US
;   CURRENT APPLICATION NUMBER: US/10/370,715B
;   CURRENT FILING DATE: 2003-02-21
;   NUMBER OF SEQ ID NOS: 742
;   SEQ ID NO 639

```

```

; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-370-715B-639
Query Match 47.4%; Score 1374.2; DB 18; Length 3443;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGTGCTCCAGACCACTTCACTGTGAGGTCCTCCCAAGAGAGTGTGTGTGTGTGTGTGT 141
Db 87 CTACTGTATCTCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 146
QY 142 TCCAGCGCCGAGCCAGAGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 201
Db 147 CGGGGCGCTCCGGGCCCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 194
QY 202 TGGCGAGCGGCTCATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 261
Db 195 TGGCGAGAGGGCTGATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 254
QY 262 ACCATCAAAACCAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 321
Db 255 ACCATCAAAACCAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 314
QY 322 CTTACGGGAGAGTGAAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 381
Db 315 CTTACGGGAGAGTGAAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 374
QY 382 TCAAGAAAGCAAAATCAATGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 441
Db 375 TCCGAAAGAGCAAAATCAATGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 434
QY 442 TCATGTCTTCAATCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 501
Db 435 TCATGTCTATCACTCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 494
QY 502 GCAAGATTGTGATTTGTATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 561
Db 495 GCAAGATCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 554
QY 562 AGCGGCAAGCGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 621
Db 555 AGCGGCAAGCGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 614
QY 622 TGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
Db 615 TGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 674
QY 682 TAGATGCAATGAGAAATCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 741
Db 675 TAGATGCAATGAGAAATCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 734
QY 742 GCAAGTCTTCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 801
Db 735 GCAAGTCTTCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 794
QY 802 GGAAGCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 861
Db 795 GGAAGCTTATCAAGAGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 854
QY 862 TGGTGCATGAGCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 921
Db 855 TGGTGCATGAGCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 914
QY 922 GTAACGGGAGCTTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 981
Db 915 GCAACGGGAGCTTACCGGAGAGCACTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 974
QY 982 TGTTAATGTGAACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1041
Db 975 TGTTAATGTGAACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1034

```

QY 1042 TCAACTGGGGTTACACCAACCGAGTGGGGGAAAGAGCCCTGCGTGAAGGGTGGGACC 1101
 DB 1035 TCAACTGGGGTTACACCAACCGAGTGGGGGAAAGAGCCCTGCGTGAAGGGTGGGACC 1094
 QY 1102 CTAGTGTGACTTTTGGCGGGGCTTCAATGGCGGACTGGTATCTGCTGCTCGGCCCC 1161
 DB 1095 CTGGCAGTGAATCTGGCCCGGCTTCAATGGCTGACTGGCTCGGCCGTTCCCGGCCCC 1154
 QY 1162 TCCGTGAGATGAGCCAGGTGTGCACTTTTCAAGCAGACGTGCGGAGAGTGA 1221
 DB 1155 TCCGTGAGATGAGCCAGGTGTGCACTTTTCAAGCAGACGTGCGGAGAGTGA 1214
 QY 1222 GCACTTACCTGGGCTGAGAGCGCAACTTCTTTAAGAAAGTCCCGAAGGAGATGA 1281
 DB 1215 GCAACACCCCTGGCTGAGAGCGCAACTTCTTTAAGAAAGTCCCGAAGGAGATGA 1274
 QY 1282 TGGCTCAAAATCTGCAAGGTGACCCGGCTGAGTACTCTTCTGCGCTTGGCAAGCA 1341
 DB 1275 TGGCCCAATCTCTCAAGTGAACCGGCTGATGAATGCCCATGCGCTTGGCAAGCA 1334
 QY 1342 GCCTTAAGCTTCCGAAAGGCACTTCTCAAGAAAGTCTCTTACTCTGCTGAGGGAGTAC 1401
 DB 1335 ACCCTCAAGCTGCGAAAGGCACTTCTCAAGAAAGTCTCTTACTCTGCTGAGGGAGTAC 1394
 QY 1402 AGGAGACCTCTCAGGAATCTGACCGGGCTGATATCTCCAGGGCAGGCTGCTGCTG 1461
 DB 1395 AGGAGACCTCTCAGGAATCTGACCGGGCTGATATCTCCAGGGCAGGCTGCTGCTG 1448
 QY 1462 TATCTCTGCTCCAGGAAAGGCACTCTTAAGAGTCTGACAGCGTGAATCTGGTACT 1521
 DB 1449 ---CGCTGCTCCAGGAAAGGCACTCTTAAGAGGCGGCAAGCGGAGTCTGGCTACT 1505
 QY 1522 ACTCTCTCCAGGAAAGGCACTCTTAAGAGGCGGCAAGCGGAGTCTGGCTACT 1581
 DB 1506 ACTCTCTCCAGGAAAGGCACTCTTAAGAGGCGGCAAGCGGAGTCTGGCTACT 1565
 QY 1582 GTGGGAGACCTCTCAGGAATCTGACAGGGCTTCAAGGGCTCTCTCCACCGGAGG 1641
 DB 1566 GTGGGAGATCCCAAGGAGCAAGAGCTCTCAGAGTTTCAAGGGCTCTCTCCACCGGAGG 1625
 QY 1642 GCAATCTCAAACTCAATGAGCAAGTCTCCGCAAGCGCTTGAAGGCACTTACCTTGA 1701
 DB 1626 GCAATCTCAAACTCAATGAGCAAGTCTCCGCAAGCGCTTGAAGGCACTTACCTTGA 1685
 QY 1702 CTTTGGCTCTCTGAGCAACTGGCTCTCTTCCATCTGCAAGCGGCGGCAAGCGGCT 1761
 DB 1686 CTTTGGCTCTCTGAGCAACTGGCTCTCTTCCATCTGCAAGCGGCGGCAAGCGGCT 1745
 QY 1762 CAGGGGCTGAGGAGAGCAAGCACTCTGCTCCGAGTCTTGAACCAATTGAGCTTGC 1821
 DB 1746 CAGGGGCTGAGGAGAGCAAGCACTCTGCTCCGAGTCTTGAACCAAGTGAATTGC 1805
 QY 1822 CTGAAGCTCTTCCGAAAGCCCACTGAAGGGCTGTGTCTGTGAACAACCTGAGGGG 1881
 DB 1806 CTGAAGCTCTTCCGAAAGCCCACTGAAGGGCTGTGTGTGTGTGAACAACCTGAGGGG 1865
 QY 1882 TTGAGCAGCTCTCTCAAGG-----GTCTGAAGCATGTGTGCAAGGATCTT 1929
 DB 1866 TTGAGGAGCCCTCTCAGAGGGCCCTGAGAGCTGCTGAGGCGCTGCGGCAAGATCTT 1925
 QY 1930 TGGGGGATAGCTGCTTCTCTGACAGACTGCAAGAGGTGACGAGCCCTTACAGAGAG 1989
 DB 1926 TGGGGGATAGCTGCTTCTCTGACAGACTGCAAGAGGTGACGAGCCCTTACAGAGAG 1985
 QY 1990 CCCTAAGAAATCTGCTCAAAAGCTGAGTGAAGAGGAGATGTGCTTATGATGGGTAG 2049
 DB 1986 CACTGAAGGATCTGCTCAAAAGCTGACCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2044
 QY 2050 GCTCTGAGAGGTTTGAAGAGAACTCTGGTGGATTTCTCC 2092
 DB 2045 GCTCTGAGATGAGCTGTGTGACCCCGAGGGGAGATGCTTC 2087

RESULT 11
 US-10-618-941-11
 ; Sequence 11, Application US/10618941
 ; Publication No. US2004019792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: MANNING, GERRARD
 ; APPLICANT: CANEPELLE, SEAN
 ; TITLE OF INVENTION: NOVEL KINASES
 ; FILE REFERENCE: 034536-0321
 ; CURRENT APPLICATION NUMBER: US/10/618,941
 ; PRIOR FILING DATE: 2003-07-15
 ; PRIOR APPLICATION NUMBER: 60/395,632
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 11
 ; LENGTH: 3463
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-618-941-11

Query Match 47.4%; Score 1374.2; DB 18; Length 3463;
 Best Local Similarity 81.6%; Pred. No. 0;
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTGCCGACCACTTCACTTGGGCTCCGACCATGAGTGGCTTAC 141
 DB 148 CTATTGATTCCTCCGCGGCTTGTCTTCACTCTGCTGCTGAGTGGCTTTCG 207
 QY 142 TCAGAGGCGGCGGACCGAGGCTCCCTGCGGCTCCGCTGCGGAGAGGCGGCGG 201
 DB 208 CGGCGGCTCCGCGGCGGCTCCCTGCGGCTCCGCTGCGGAGAGTGA-----GCCCGGCGG 255
 QY 202 TGGCGGAGCGGCTCATCAAGTCCCTTAACTCTGATGAAGAGCAGGCGGTGAAGCGG 261
 DB 256 TGGCGGAGAGGCTGATCAATGCTCCGAGGCTTATGAAGAGCAGGCGGTGAAGCGG 315
 QY 262 ACCATCAAAACAACTTGGGCGACCGCTTGAAGTCTTGAAGAGCGTGGGAGGCA 321
 DB 316 ACCATCAAAACAACTTGGGCGACCGCTTGAAGTCTTGAAGAGCGTGGGAGGCA 375
 QY 322 CTTACGGGAGAGTGAAGAGGCAAGAGGCTTGGGCGCTCTGTGAGCCATCAAGTCCA 381
 DB 376 CTTACGGGAGAGTGAAGAGGCGGAGGCTTGGGCGCTCTGTGAGCCATCAAGTCCA 435
 QY 382 TCAGAAAGCAAAATCAAGATGAGCAAGGATCTGCTGCAATTAAGAGGAGATTGAGA 441
 DB 436 TCAGAAAGCAAAATCAAGATGAGCAAGGATCTGCTGCAATTAAGAGGAGATTGAGA 495
 QY 442 TCATGCTTCACTCAACCAACCCCAATCAATTCATCAATCAATCAATCAATCAATCA 501
 DB 496 TCATGCTATCACTCAACCAACCCCAATCAATTCATCAATCAATCAATCAATCAATCA 555
 QY 502 GCAAGATGTGATGTCTATGAAGATGAGCGGAGGCGGATCTGTATGATTAATCAAGT 561
 DB 556 GCAAGATGTGATGTCTATGAAGATGAGCGGAGGCGGATCTGTATGATTAATCAAGT 615
 QY 562 AGCGGCAAGCGGCTGAGTGAAGCGGAGCGGAGGATTTCTTCCGACAGATGTGTGCGCC 621
 DB 616 AGCGGCAAGCGGCTGAGTGAAGCGGAGCGGAGGATTTCTTCCGACAGATGTGTGCGCC 675
 QY 622 TGCATTAATGCAACCAAGGAGTGTCAACCGAGATCTCAAGTCTGAGAGAAATCTCTTC 681
 DB 676 TGCATTAATGCAACCAAGGAGTGTCAACCGAGATCTCAAGTCTGAGAGAAATCTCTTC 735
 QY 682 TAAATGCAATGAGAAATCAAGATGTGATGCTTGGCTCTCAACCTGTACCAAGAG 741
 DB 736 TGAATGCAATGAGAAATCAAGATGTGATGCTTGGCTCTCAACCTGTACCAAGAG 795
 QY 742 GCAAGTTCCTCAAGAGTCTGAGGAGCGCTCTCAAGCTGAGCTGAGATGTAAGAG 801
 DB 796 GCAAGTTCCTCAAGAGTCTGAGGAGCGCTCTCAAGCTGAGCTGAGATGTAAGAG 855

QY 802 GGAAGCCCTATGTTGGGCCCCAGAGGTGAGCAGCTGCTCTGAGGCGTTCTCTGTAATCC 861
DB 856 GGAAGCCCTACACAGGCCCAAGAGGTGAGCAGCTGCTCTGAGGCGTTCTCTGTAATCC 915
QY 862 TGTGTCATGTCACATGCTCCCTTTTGAACGGGCGAGTCAATAAACATGCTGAAGCAATCA 921
DB 916 TGTGTCATGTCACATGCTCCCTTTTGAACGGGCGAGTCAATAAACATGCTGAAGCAATCA 975
QY 922 GTAAAGGGGCTTACCGTGAAGCGCGCCAAAGCGTCCAGTCTGTGAGCTGATCCGGTGGC 981
DB 976 GCAACGGGGCTTACCGGAGGACCTTAACCTCTGATGCTGTGAGCTGATCCGGTGGC 1035
QY 982 TGTAAATGTGAACCCCAACCGGTGGGCGCACTGAGAGATGTAAGCAATGTTGAGG 1041
DB 1036 TGTAAATGTGAACCCCAACCGGTGGGCGCACTGAGAGATGTAAGCAATGTTGAGG 1095
QY 1042 TCAACTGGGGTTTACCAACCGAGTCCGGGGAACAGAAACCTTGCGTGAAGGTGGCAAC 1101
DB 1096 TCAACTGGGGTTTACCGCAACCGAGTGGGAGAGAGAGAGCTCCGCAATGAGGGTGGCAAC 1155
QY 1102 CTAGTGTGACTTGTGGCGGGGCTTCAGTGGGCGCACTGGTTTACGTGCTCTCGGGCCCC 1161
DB 1156 CTGCAAGTGAATCTGCGCGGCTTCAGTGGGCGCACTGGTTTACGTGCTCTCGGGCCCC 1215
QY 1162 TCCTGAGAAATGAGCAAGGTGTGCACTTCTTCAAGCAGCAGTGGCGGAGAGTGGAA 1221
DB 1216 TCCTGAGAAATGAGCAAGGTGTGCACTTCTTCAAGCAGCAGTGGCGGAGAGTGGAA 1275
QY 1222 GCACTGTAATCTGGGCTGAGAGCGGCAATCTTCTTGAAGAGTCCGAAAGAGATGACA 1281
DB 1276 GCACTGTAATCTGGGCTGAGAGCGGCAATCTTCTTGAAGAGTCCGAAAGAGATGACA 1335
QY 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCCTCTTCTCGCCCTGGCAAGACA 1341
DB 1336 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCCTCTTCTCGCCCTGGCAAGACA 1395
QY 1342 GCCTTAAGCTTCCGAAAGGCAATCTTCAAGAAAGTCTCTTCACTGCTCAAGGAGGTAC 1401
DB 1396 ACCTCAAGCTTCCGAAAGGCAATCTTCAAGAAAGTCTCTTCAAGGAGGTAC 1455
QY 1402 AGAGAGACCTTCAGAGAACTGACCGGTGCTGATCTCAAGGAGAGCTGCTGCTGCTG 1461
DB 1456 AGAGAGACCTTCAGAGAACTGACCGGTGCTGATCTCAAGGAGAGCTGCTGCTGCTG 1509
QY 1462 TATCTCTGCTCCCAAGAGAGGCAATCTTGAAGAGTCTCAAGAGCTGAGTCTGCTACT 1521
DB 1510 ---CCCTGCTCCCAAGAGAGGCAATCTTCAAGAGAGCTGAGTCTGCTACT 1566
QY 1522 ACTCTCTCCAGAGAGGAGGAGTCTGAGAGAACTCTTGAAGAGGAGAGTGTGTTTGA 1581
DB 1567 ACTCTCTCCAGAGAGGAGGAGTCTTGAAGAGGAGTCTTGAAGAGGAGGAGGAGTGTG 1626
QY 1582 GTGGGAGACCCGTGAGAGAGAGTCTCAAGGCTTCAAGGCTCTCTCTCAAGGAGAG 1641
DB 1627 GTGGGAGATCCCAAGAGAGAGGCTCTCAAGGCTTCAAGGCTCTCTCTCAAGGAGAG 1686
QY 1642 GCATTTCAAACTCAATGAGCAAGTCTCTCGGCAAGAGCTTGAAGAGCACTAACCTTGA 1701
DB 1687 GCATTTCAAACTCAATGAGCAAGTCTCTCGGCAAGAGCTTGAAGAGCACTAACCTTGA 1746
QY 1702 CTTTGGCTCCGTGAGCAACTGGGCTCTCTCCAGCTGAGAGCGCGGCGCAAGCGGCT 1761
DB 1747 CTTTGGCTCCGTGAGCAACTGGGCTCTCTCCAGCTGAGAGCGCGGCGCAAGCGGCT 1806
QY 1762 CAGGGGCTGTGAGTGAAGAGAGCTGTCTCTCGAGTCTTGAACCAATGAGACTTGC 1821
DB 1807 CAGGGGCTGTGAGTGAAGAGAGCTGTCTCTCGAGTCTTGAACCAAGTGAAGCTTGC 1866
QY 1822 CTGAAGCTTCTCCGAAAGCCCACTGAAGGGGCTGTGAGTCTGTGAGCAACTGAGGGGCG 1881
DB 1867 CTGAAGCTTCTCCGAAAGCCCACTGAGGGGCTGTGAGTCTGTGAGCAACTGAGGGGCG 1926

QY 1882 TTGAGAGCTCTCCCTCAAGAG-----GTCTGAAGCAGATGTGTGAGAGATCT 1929
DB 1927 TTGAGAGCCCCCTCAAGAGGCCCTGGAAGCTCTGAGAGCGCTGCGGAGGATCTT 1986
QY 1930 TGGGGAGTAGCTCTTCTTCTGACAGACTGCGCAAGAGGTGACTGAGCCTCAAGACAAG 1989
DB 1987 TGGGGAGTAGCTCTTCTTCTGACAGACTGCGCAAGAGGTGACTGAGCCTCAAGACAAG 2046
QY 1990 CCTGAGAAATCTGCTCAAGAGCTGAGAGAGAGAGATGTGCTCTGATGTGGGTAG 2049
DB 2047 CACTGAGAGGTCTGCTCAAGAGCTGAGAGAGAGATGTGCTCTGATGTGGGTAG 2105
QY 2050 GCTCTGAGAGGTTTGAAGAGAACTCTGGGCTGAGTCTCC 2092
DB 2106 GCTCTGAGATGACAGCTGTTGACCCCGAGGGAGAGATGCTTC 2148

RESULT 12
US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids a
; TITLE OF INVENTION: PolyPeptides
; FILE REFERENCE: 803.1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 47.3%; Score 1372.2; DB 17; Length 2501;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTGCCAGACCACTTCACCTGCGGCTCCCGACCATGAGTGGTGGCTTAC 141
DB 65 CTATGATTCCTCCCTGCGGCTTGTCTCACTCTGCTGCGCATGAGTCTGCTGTTTCG 124
QY 142 TCCAGGCGCCGAGCAGGCTCCCTCGGCTCCCGCTGCGCTCGAGAGCGCCCGGCGC 201
DB 125 CGGCGGCTCCGCGCCCACTCTCTCGGCTCGAGAGCTA-----GCCGAGCTGC 172
QY 202 TGGCGGAGGGGCTCATGAGTCTGCTTAACTCTGATGAGAGAGCAGGCGTGAAGGCG 261
DB 173 TGGCGGAGGGGCTCATGAGTCTGCGCAAGCCCTTAATGAGAGAGAGCGGTGAAGGCG 232
QY 262 ACCATCAACAACACACTCTGGGAGCAGCTTACGAGTCTCTGAGAGAGCTGGGCAAGGCA 321
DB 233 ACCACACAAGACACACTCTGGGAGCAGCTTACGAGTCTCTGAGAGAGCTGGGCAAGGCA 292
QY 322 CCTACGGGAGAGTGAAGAGAGAGAGAGAGTCTGGGCGGCTGTGAGCAATCAAGTCA 381
DB 293 CCTACGGGAGAGTGAAGAGAGAGAGAGAGTCTGGGCGGCTGTGAGCAATCAAGTCA 352
QY 382 TCAGAGAAAGCAAAATCAAAAGATGAGAGATCTGTGACATTCGAGAGAGAGATTGAG 441

Db 353 TCCGGAAGACAAATCAAGATGAGCAAGATCTGATGCACATACGAGGGAGATTGGA 412
Qy 442 TCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAAGATGCA 501
Db 413 TCATGTCACTCAACCAACCCCAATCATTTGCCATCATGAAAGTTTGAAGATGCA 472
Qy 502 GCAAGATTGTGATTTGTCAATGATGACCGAGCGGATCTGTATGATTAACATCAGTG 561
Db 473 GCAAGATTGTGATTTGTCAATGATGACCGAGCGGATCTGTATGATTAACATCAGTG 532
Qy 562 AGCGGCAACGCTGATGAGCGGAGCGGAGCGGATCTGTATGATTAACATCAGTG 621
Db 533 AGCGGCAACGCTGATGAGCGGAGCGGAGCGGATCTGTATGATTAACATCAGTG 592
Qy 622 TGCACTATCTGCAACCAAGAGGATCTGTACCGAGATCTAGTGGAAACATCTTCC 681
Db 593 TGCACTATCTGCAACCAAGAGGATCTGTACCGAGATCTAGTGGAAACATCTTCC 652
Qy 682 TAGATGCAATGAGAAACATCAAGATGCTGATCTTGGCTCTGCAACCTGACCAAG 741
Db 653 TAGATGCAATGAGAAACATCAAGATGCTGATCTTGGCTCTGCAACCTGACCAAG 712
Qy 742 GCAAGTCTTCAACGCTTCTGTGAGCGGCTCTCTACGCTGCTGAGATGATCAAG 801
Db 713 GCAAGTCTTCAACGCTTCTGTGAGCGGCTCTCTACGCTGCTGAGATGATCAAG 772
Qy 802 GGAAGCCCTATGTGGGCGGAGAGGATGACAGCTGCTGTGGCGTTCTCTGATCACC 861
Db 773 GGAAGCCCTATGTGGGCGGAGAGGATGACAGCTGCTGTGGCGTTCTCTGATCACC 832
Qy 862 TGGTGCATGAGCAACATGACCGGCGAGATCAATTAACATCTGTGAAGAAATCA 921
Db 833 TGGTGCATGAGCAACATGACCGGCGAGATCAATTAACATCTGTGAAGAAATCA 892
Qy 922 GTAAAGGAGCTTACCGTGAAGCGGCGCAAGCGTCCAGTCTGTGAGCTGATCCGATGC 981
Db 893 GTAAAGGAGCTTACCGGAGCGGCGCAATTAACCGCTGTGAGCTGATCCGATGC 952
Qy 982 TGTATTAATGTGAACCCCAACCGCTGCGGCGCACTGAGAGATGTAACCATGATGTTGG 1041
Db 953 TGTATTAATGTGAACCCCAACCGCTGCGGCGCACTGAGAGATGTAACCATGATGTTGG 1012
Qy 1042 TCAACTGGGGTTACCAACCGGAGTCCGGGAGCAAGAACCTGTCGTAAGGATGAGCAAC 1101
Db 1013 TCAACTGGGGTTACCAACCGGAGTCCGGGAGCAAGAACCTGTCGTAAGGATGAGCAAC 1072
Qy 1102 CTAGTGTGATCTTGGCGGCGCTCTCAATGCGGAGCTGCTGCTCTGCGGCCCC 1161
Db 1073 CTGCAATGATCTGCGCGGCGCTCTCAATGCGGAGCTGCTGCTCTGCGGCCCC 1132
Qy 1162 TCTGTGAAGATGAGCAAGGATGTCAGCTTCTTAAGCAGACGTCGCGGAGATGAG 1221
Db 1133 TCTGTGAAGATGAGCAAGGATGTCAGCTTCTTAAGCAGACGTCGCGGAGATGAG 1192
Qy 1222 GCACTGATACCTGGGCTGAGAGGAGCAATCTTGAAGAAATCCCGAAAGAGATGACA 1281
Db 1193 GCACTGATACCTGGGCTGAGAGGAGCAATCTTGAAGAAATCCCGAAAGAGATGACA 1252
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTTCTGCGCTGCAAGACA 1341
Db 1253 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATACCTTCTGCGCTGCAAGACA 1312
Qy 1342 GCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGAGTCTCTGATCTGCTGAGAGGATG 1401
Db 1313 ACCTCAAGTGCCTCAAGAGGCAATCTCAAGAAAGAGTCTCTGATCTGCTGAGAGGATG 1372
Qy 1402 AGAGAGACCTCTCAAGAACTCAAGACCGGCTGATACCTCAAGAGGAGGATGCTGCTG 1461
Db 1373 AGAGAGACCTCTCAAGAACTCAAGACCGGCTGATACCTCAAGAGGAGGATGCTGCTG 1426
Qy 1462 TATCCCTGCTCCCAAGAGGAGGATCTTAAAGATCTGCAAGAGGATGATGATGATG 1521
Db 1427 TATCCCTGCTCCCAAGAGGAGGATCTTAAAGATGCTGCAAGAGGATGATGATGATG 1483

Qy 1522 ACTCTCTCCAGAGCCCAAGGATCTGAGGAACTTTAGACGCCATGATGTTTGTGA 1581
Db 1484 ACTCTCTCCAGAGCCCAAGGATCTGAGGAACTTTAGACGCCATGATGTTTGTGA 1543
Qy 1582 GTGGGAGCCCGTGGAGCAAGAGTCTCAGAGGCTTCAAGGCTCTCTCCACAGGAG 1641
Db 1544 GTGGGAGTCCCGAGAGCAAGAGTCTCAGAGGCTTCAAGGCTCTCTCCACAGGAG 1603
Qy 1642 GCATTTCAAACTCAATGAGCAAGTCTCCGCAACAGCTTGAAGAGCACTAACCCCTG 1701
Db 1604 GCATTTCAAACTCAATGAGCAAGTCTCCGCAACAGCTTGAAGAGCACTAACCCCTG 1663
Qy 1702 CCTTGGCTCTCCAGAGCAAGTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCT 1761
Db 1664 CCTTGGCTCTCCAGAGCAAGTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCT 1723
Qy 1762 CAGGAGCTGAGAGGAGCAAGATCTGCTCCGAGTCTTGAACCAATGAGATGAG 1821
Db 1724 CAGGAGCTGAGAGGAGCAAGATCTGCTCCGAGTCTTGAACCAATGAGATGAG 1783
Qy 1822 CTGAACGCTCTTCCGAAACCCCACTGAGGAGCTGAGTCTGAGCAACCTGAGGAG 1881
Db 1784 CTGAACGCTCTTCCGAAACCCCACTGAGGAGCTGAGTCTGAGCAACCTGAGGAG 1843
Qy 1882 TTGAGCAGCTCTCTCAAGG-----GTCTGAAGCATGCTGAGGATGCT 1929
Db 1844 TTGAGCAGCTCTCTCAAGG-----GTCTGAAGCATGCTGAGGATGCT 1903
Qy 1930 TGGGGGATGATGCTTCTTCTGACAGACGAGGAGGATGATGATGATGATGATGATG 1989
Db 1904 TGGGGGATGATGCTTCTTCTGACAGACGAGGAGGATGATGATGATGATGATGATG 1963
Qy 1990 CCTGAGGATGATGCTTCTTCTGACAGACGAGGAGGATGATGATGATGATGATGATG 2049
Db 1964 CACTGAGGATGATGCTTCTTCTGACAGACGAGGAGGATGATGATGATGATGATGATG 2022
Qy 2050 GCTCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2090
Db 2023 GCTCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2063

RESULT 13
US-10-311-034-38
Sequence 38, Application US/10311034
Publication No. US2004023242A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
INVENTOR: YUB, Henry
APPLICANT: LAL, Preeti
APPLICANT: BANDMAN, Olga
APPLICANT: BOROMSKY, Mark L.
APPLICANT: AU-YOUNG, Janice
INVENTOR: LU, Van
APPLICANT: GANDHI, Ameena R.
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: CHAMLA, Narinder K.
INVENTOR: YAO, Monique G.
APPLICANT: GREENWALD, Sara R.
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: KEARNEY, Liam
APPLICANT: BURFORD, Neil
INVENTOR: NGUYEN, Daniel B.
APPLICANT: TANG, Y. Tom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: HE, Ann
APPLICANT: THORNTON, Michael
APPLICANT: HAPALIA, April
APPLICANT: ARVIZU, Chandra S.
APPLICANT: GURURAMAN, Rajagopal
INVENTOR: LO, Terence P.

APPLICANT: KEAH, Farrah A.
 APPLICANT: RECIPON, Shirley A.
 APPLICANT: AZIMZAI, Yalda
 APPLICANT: POLICKY, Jennifer L.
 APPLICANT: DING, Li
 APPLICANT: GREYER, Megan
 APPLICANT: ELIOTT, Vicki S.
 APPLICANT: THANGAVELU, Kavitha
 APPLICANT: BATRA, Sajeew
 APPLICANT: ISON, Craig H.
 TITLE OF INVENTION: HUMAN KINASES
 FILE REFERENCE: PI-0125 PCT
 CURRENT APPLICATION NUMBER: US/10/311,034
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
 60/228,056
 PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-07-25
 NUMBER OF SEQ ID NOS: 52
 SOFTWARE: PERL Program
 SEQ ID NO 38
 LENGTH: 3360
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CH1
 US-10-311-034-38

Query Match 47.2%; Score 1370.6; DB 17; Length 3360;
 Best Local Similarity 81.5%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 339; Indels 35; Gaps 4;

QY 82 CTACTGCTGCCGACCCCACTCACTGCGGCTCCCCGACCATGAGTGGTGGCTTAC 141
 DB 39 CTACTGATTCCTCCGCGCCCTTCTGCTCACTGCTGCGCATGAGTGGTGGTTCG 98
 QY 142 TCCAGGCGCCGAGCAGGCTCCTGCGGCTCCGCGCTGCGGCTCGAGAGCGCCGCGC 201
 DB 99 CGCGGCGCTCCGCGCCCACTCTCTCGGCG-----CAGAGCTAGCCGCGCGC 145
 QY 202 TGGCGGAGCGGCTCATCAAGTCCGCTAAACCTCTGATGAGAGAGAGCGGCTGAGCGGC 261
 DB 146 TGGGGAAGGCTATCAAGTCCGCGCAAGCCCTTAAGAGAGAGCGGCTGAGAGCGGC 205
 QY 262 ACCATCAAAACAACAACCTGCGGCAAGCTTCAAGTTCCTGAGAGAGCGTGGGCAAGGCA 321
 DB 206 ACCACCAACAACAACCTGCGGCAAGCTTCAAGTTCCTGAGAGAGCGTGGGCAAGGCA 265
 QY 322 CTTACGGGAGAGTGAAGAGGCAAGAGAGCTGGGGCGTCTGGTGGCATCAAGTCCA 381
 DB 266 CTTACGGGAGAGTGAAGAGGCGGAGAGCTGGGGCGCTGGTGGCATCAAGTCCA 325
 QY 382 TCAGAGAAAGCAAAATCAAGATGAGAGATCTGCTGACATACGAGAGAGATTGAGA 441
 DB 326 TCCGGAAGAGCAAAATCAAGATGAGAGATCTGATGAGCATACGAGAGAGATTGAGA 385
 QY 442 TCATGCTTCACTCAACCAACCCCAATCATTTGCCATTCATGAAGTTTGAAGATTGCA 501
 DB 386 TCATGCTATCGCTCAACCAACCTCAATCATTTGCCATTCATGAAGTTTGAAGATTGCA 445
 QY 502 GCAAGATTGATGTGATGAGATGAGATGAGAGAGAGCTGATGATGATTAATTAACATTCAGT 561
 DB 446 GCAAGATTGATGTGATGAGATGAGATGAGAGAGAGCTGATGATGATTAATTAACATTCAGT 505
 QY 562 AGCGGCAACGCTGAGTGAAGCGGAGCGCAAGCAATTTCTTCCGACAGATCTGTCTGGCC 621
 DB 506 AGCGGCAACGCTGAGTGAAGCGGAGCGCAAGCAATTTCTTCCGAGAGATCTGTCTGGCG 565
 QY 622 TGCATCTAGTCCACCAAGAGCGGATCTTCAACGAGATCTCAAGTGGAGAAATCTCTTC 681
 DB 566 TGCATCTAGTCCACCAAGAGCGGATCTTCAACGAGATCTCAAGTGGAGAAATCTCTTC 625

QY 682 TAGATGCCAATGGAACATCAAGATTCCTGACTTTGGCCCTCTCCACCTGATCAACAAG 741
 DB 626 TGGGTGCCAATGGAATATCAAGATTCCTGACTTTGGCCCTCTCCACCTGATCAACAAG 685
 QY 742 GCAAGTCCCTCAAGCTTCTGAGGAGCCCTCTCAAGGCTTGGCTGAGATGATCAAG 801
 DB 686 GCAAGTCCCTCAAGCTTCTGAGGAGCCCTCTCAAGGCTTGGCTGAGATGATGATG 745
 QY 802 GGAAGCCCTATGTTGGGCGCAGAGGTGAGACAGTGTCTGGGCGTCTCTGATACATCC 861
 DB 746 GGAAGCCCTATGAGGCGCAGAGGTGAGACAGTGTCTCTGGGCGTCTCTGATACATCC 805
 QY 862 TGGTCATGACGACCAATGCTTTGACGAGGAGATCAATAAACAATGTTGAGAAATCA 921
 DB 806 TGGTCATGACCAATGCTTTGATGAGGAGATGACATAAGATCTTATGTAACAGATCA 865
 QY 922 GTAAAGGCGCTTACCGTGAAGCGCGCCCAAGCGTCCGATGCTGGGCTGATCGGATGCG 981
 DB 866 GCAAGGCGGCTTACCGGAGGACATTAACCTCTGATGCTTGGGCTGATTCGGATGCG 925
 QY 982 TGTATATGTTGAACCCCAACCGGTCGAGCACTGAGAGATGATGATGATGATGATG 1041
 DB 926 TGTATATGTTGAACCCCAACCGGTCGAGCACTGAGAGATGATGATGATGATGATG 985
 QY 1042 TCAACTGAGGTTTACCAACCGGATCGGAGAACAGAAAGCGTGTGAGGTTGAGCA 1101
 DB 986 TCAACTGAGGTTTACCAACCGGATCGGAGAACAGAAAGCGTGTGAGGTTGAGCA 1045
 QY 1102 CTACTGATGATTTGGCGGCGCTTCATGAGCGGATGATGATGATGATGATGATG 1161
 DB 1046 CTGCACTGATGATTTGGCGGCGCTTCATGAGCGGATGATGATGATGATGATG 1105
 QY 1162 TCCGAGAGATGAGAGCAAGGTGAGAGCTTCTTCAAGAGACAGATGCGGAGAGTGA 1221
 DB 1106 TCCGAGAGATGAGAGCAAGGTGAGAGCTTCTTCAAGAGACAGATGCGGAGAGTGA 1165
 QY 1222 GCACTGATGATGAGGCTGAGAGCGGCAATTTCTTAAAGATGCCGAAGAGATGCA 1281
 DB 1166 GCACTGATGATGAGGCTGAGAGCGGCAATTTCTTAAAGATGCCGAAGAGATGCA 1225
 QY 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTTCTTGGCTTGGCAAGAGA 1341
 DB 1226 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTTCTTGGCTTGGCAAGAGA 1285
 QY 1342 GCCTTAAGCTTCCGAAGAGATCTTCAAGAAAGTCTCTCAAGCTTCCGAGGAGATG 1401
 DB 1286 ACCTCAAGCTTCCGAAGAGATCTTCAAGAAAGTCTCTCAAGCTTCCGAGGAGATG 1345
 QY 1402 AGAGAGACCTTCAAGAGATCTGAGACCGGCTGATGATCTTCCAGGAGAGCTTCTG 1461
 DB 1346 AGAGAGACCTTCCGAGATCTGAGACCGGCTGATGATCTTCCAGGAGAGCTTCTG 1399
 QY 1462 TATCTGCTTCCGAAGAGAGATCTTAAAGATCTTCAAGAGAGATGATGATGATG 1521
 DB 1400 ---CCCTGCTTCCGAAGAGAGATCTTCAAGAGAGATGATGATGATGATG 1456
 QY 1522 ACTCTCTTCCAGAGACCTTCAAGAGATCTTAAAGATCTTCAAGAGAGATGATGATG 1581
 DB 1457 ACTCTCTTCCAGAGACCTTCAAGAGATCTTAAAGATCTTCAAGAGAGATGATGATG 1516
 QY 1582 GTGGGAGACCTTCAAGAGATCTTCAAGAGATCTTCAAGAGATGATGATGATGATG 1641
 DB 1517 GTGGGAGATCCCAAGAGATCTTCAAGAGATCTTCAAGAGATGATGATGATGATG 1576
 QY 1642 GCATTTCAACTCAATGAGCAAGTTCTCCGCAAGCTTGAAGAGCACTAACCTTAC 1701
 DB 1577 GCATTTCAACTCAATGAGCAAGTTCTCCGCAAGCTTGAAGAGCACTAACCTTAC 1636
 QY 1702 CTTTGGCTCCCTGAGCAACATGAGCTTCCCTTCCATCTGAGAGAGCGGCGGCGGCT 1761
 DB 1637 CTTTGGCTCCCTGAGCAACATGAGCTTCCCTTCCATCTGAGAGAGCGGCGGCGGCT 1636
 QY 1762 CAGGGGCTGATGAGAGAGATCTTCTTCCAGAGCTTCTTGAACCAATGATGATG 1821

Db 1697 CAGGGGCTGTGAGGAGGAGGAGCATTCCTGTCTCTGAGTCTCTTTGACCAAGTGGACTTTGC 1756
Qy 1822 CTGAAGCTCTTCCGAAACCCTACTGAGGGGCTGTGTCTGTGACAACCTTGAGGGGCGC 1881
Db 1757 CTGAACGGCTCCAGAGAGCCCTACTGCGGGGCTGTGTCTGTGTGACAACTTCAAGGGGC 1816
Qy 1882 TTGAGCAGCCTCCCTCAAGAG-----GTCTGAAGCGATGGTGGAGGAAATCCT 1929
Db 1817 TTGAGAGAGCCCTCTCAAGAGGCGCTGGAAGCTCTGAGGCGCTGGGCGGAGATCCTT 1876
Qy 1930 TGGGGGATAGCTGTCTTCTCTGACAGACTGCCAAGAGTGTACTGACCTTACAGACAG 1989
Db 1877 TGGGGGACAGCTGCTTTTCCCTGACAGACTGCCAAGAGTGTACTGACCTTACAGACAG 1936
Qy 1990 CCTTGAAGATCTGCTCAAGCTCAGCTGAGAGAGGAGATGTCCTTATGATGAGGATG 2049
Db 1937 CACTGAGGGTCTGCTCAAGCTCAGCTGAGTGTGAGTGTGAGCATTCGCCAG-CCGGTCTAG 1995
Qy 2050 GCTCTGAGAGGCTTTGACAGAGAACCTCGGGTGGATTCCTCC 2092
Db 1996 GCTCTGAGATGACAGCTGCTGTGACCCCGAGGGAGATGCTTTC 2038

RESULT 14
US-09-963-159-3
; Sequence 3, Application US/09963159
; Patent No. US20020077312A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700. A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-5001
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-963-159-3

Query Match 46.9%; Score 1360.2; DB 9; Length 1884;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

Qy 169 CTTCCGCGCTGGGCTGAGAGAGCGCGCGCTGGGCGGACGGGCTCATCAAGTCGCTA 228
Db 35 CCACTCCCTCGGCGGAGAGCTAGCCGCGCTGGCGGAAGGGCTGATCAAGTCGCCA 94
Qy 229 AACCTTGATGAAGAACAGAGCGGTGAGCGGACATCAACAACAACCTGGCGGACC 288
Db 95 AGCCCTGAAGAAGAACAGAGCGGTGAGCGGACATCAACAACAACCTGGCGGACC 154
Qy 289 GCTACGAGTTCCTGAGAGACGCTGGGCAAGGCGACTTA CGGAGAGTGAAGAGCGACAG 348
Db 155 GCTACGAGTTCCTGAGAGACGCTGGGCAAGGCGACTTA CGGAGAGTGAAGAGCGCGG 214
Qy 349 AGAGCTCGGGGCGTGTGTGGCCATCAAGTCCATCAAGAAAGCAAAATCAAAATGAGC 408
Db 215 AGAGCTCGGGGCGTGTGTGGCCATCAAGTCCATCAAGTCCGAAAGCAAAATCAAAATGAGC 274
Qy 409 AGGATCTGCTGCAATACGAGAGGAGATTGATCATGTCTTCACTCAACCAACCCGACA 468
Db 275 AAGATCTGATGCAATACGAGAGGAGATTGATCATGTCTTCACTCAACCAACCCGACA 334
Qy 469 TCATTGCTATCATGAAGTGTGGAAGTAAGCAAGCAAGATTGTGATGATGAGATG 528
Db 335 TCATTGCTATCATGAAGTGTGGAAGTAAGCAAGCAAGATTGTGATGATGAGATG 394

Qy 529 CCAGCCGAGCGATCTGTATGATTATCACTGAGAGCGGCCACCGCTGATGATGAGCGGACG 588
Db 395 CCAGCCGAGCGCGACTTTATGATCAATCAACGAGCGGCGAGAGCTGATGAGCGGCAAG 454
Qy 589 CAGGCAATTTCTTCCGACAGATGATGTCTGCGCTGCACTAGTGCACAGAACGGATGCG 648
Db 455 CTAGGCAATTTCTTCCGCAATGCTCTCTGCGGCACTATTTGGCACTAGAACAGATGG 514
Qy 649 TTCAACGAGATCTCAAGCTGAAAACATCTTCTAGATGCCAATGAAATCAAGATTG 708
Db 515 TTCAACGAGATCTCAAGCTGAAAACATCTTCTAGATGCCAATGAAATCAAGATTG 574
Qy 709 CTGACTTGGGCTCTCCAACTGTACACAAAGGCAAGTCTTCAACGCTTGTGAGGA 768
Db 575 CTGACTTGGGCTCTCCAACTGTACACAAAGGCAAGTCTTGTGAGGA 634
Qy 769 GCGCTCTCTAGCGCTCGCTGAGATAGTCAACGAGGAGCCCTTATGTTGGGCGCAGAGTGG 828
Db 635 GCGCGCTCTATGCTGCGCCAGAGATTTCAATGGAAAGCTTACACAGGCTCCAGAGTGG 694
Qy 829 ACAAGTGTCTGTGGGCGTTCTCTGTACATCTCTGTGATGAGCAACGATCCCTTTGACG 888
Db 695 ACAAGTGTCTGTGGGCGTTCTCTGTACATCTCTGTGATGAGCAACGATCCCTTTGATG 754
Qy 889 GGCAGGATCAATAAACACTGTGAGAGAAATCACTTAAACGGGGCTTAACTGTAGCGGCCCA 948
Db 755 GGCAGGATCAATAAAGATCTGTAGTAACAGATCAAGCAACGGGGCTTAACTGTAGCGGCCCA 814
Qy 949 AGCGTCCGATGCGCTGTGCGCTGATCCGATGCGGTGTATATGATGAACCCGACCGTGGG 1008
Db 815 AACCTCTGATGCTGTGTGCGCTGATCCGATGCGGTGTATATGATGAACCCGACCGGCG 874
Qy 1009 CCAACTGAGAGATGATGACCAATGATGATGATGATGATGATGATGATGATGATGATG 1068
Db 875 CCAACTGAGAGATGATGACCAATGATGATGATGATGATGATGATGATGATGATGATG 934
Qy 1069 GGAACAAGAAAGCGCTGCGGAGAGGAGGAGCACTTATGATGATGATGATGATGATGATG 1128
Db 935 GGAACAAGAAAGCGCTGCGGAGAGGAGGAGCACTTATGATGATGATGATGATGATGATG 994
Qy 1129 TGGCGGACTGTTACGTGCTCTCGCGGCGCCCTCTCGAGAAATGAGCCAGGTGTGCA 1188
Db 995 TGGCTGACTGCTCTCGCGGCGTTCTCTCGCGGCGCCCTCTCGAGAAATGAGCCAGGTGTGCA 1054
Qy 1189 GCTTCTTCAAGCAGCAGTGTGCGGAGGAGGAGCACTTATGATGATGATGATGATGATG 1248
Db 1055 GCTTCTTCAAGCAGCAGTGTGCGGAGGAGGAGCACTTATGATGATGATGATGATGATG 1114
Qy 1249 ATTCTTTAAGAGTCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1308
Db 1115 ATTCTTTAAGAGTCCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1174
Qy 1309 CTGAGGATACCTCTTCTCGCGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1368
Db 1175 CTGATGACATGCGCCATGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1234
Qy 1369 AGAAAAAGTCTTACTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1428
Db 1235 AGAAAAAGTCTTACTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1294
Qy 1429 TGCCTGATACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1488
Db 1295 TGCCTGAGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1345
Qy 1489 TTAAAGATCTGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1548
Db 1346 TCAAGAGAGCCCGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1405
Qy 1549 GGAAGCTTTTGAAGCGCAGT 1608
Db 1406 GGAAGCTTTTGAAGCGCAGT 1465
Qy 1609 CACAGGCTTACAGGAGGCTCTCTCTCCACCGCAAGGGCAATTCTCAAACTCATGCAAGTTCT 1668

QY	1069	GGGAAACAGAAAGCCCTGGCTGTAGAGGGGTGGGACACCTTAAGTGTGATCTTTTGGCCGGGCTTCCA	1128
Db	935	GAGAGCAGAGAGCTTCCGATGATGGGTGGGACACTTGGCAGTGAATCTTCCCGCTTCCA	994
QY	1129	TGGCGAATGTGTAGTGTGCTTCTGCGCCCTCTCTTGGAGATATGAACCAAGTGTGTCA	1188
Db	995	TGGCTGAATGTGCTCCGGGGTTCCTTCCGCCCCCTCTTGGAGATATGGGCTCAAGTGTGTCA	1054
QY	1189	GCTTCTTACACGACGATGCTCCCGGAGGTGGAAAGCATGTATACCTGGGCTTGGAGCGGCAAC	1248
Db	1055	GCTTCTTACACAGATATCACCTGGTGGGGAGAACACACCCTGGCTCTGGAGCGCTCAGC	1114
QY	1249	ATTCTCTTAAGAAGTCCCGAAAGGAAATGACATGGCTCAAAATATGCAAGTGAACCCGG	1308
Db	1115	ATTGGCTCAAGAAATCCCGCAAGAGATATGACATGGCCAGTCTTCTTCAAGTGAACAGG	1174
QY	1309	CTGAGATATACCTTCTTCCGCCCTGGCAAGACAGCTTAAAGCTTCCGAAAGCATTTCTCA	1368
Db	1175	CTGATGACATATGCCCATATGCTCTGGCAAGACACACTCAAGCTGGCAAAAGGCAATTTCTCA	1234
QY	1369	AGAAAGAAATCCTCTTAACTCTGTACAGGGAGGTACAGAGAGACCTCTCAGAACTTCAAGCCGG	1428
Db	1235	AGAAAGAAAGTGTACGCTCTGTACAGAGGGGTATACAGAGAGACCTCTCCGAGCTCAGCCCAA	1294
QY	1429	TGCGCTGATATCCAGGGGAGGCTGTGCCCCGTGCTGATCCCTGTCTCCCAAGAGAAAGCATTC	1488
Db	1295	TCCCTTGGAGCCCAAGGCGAGGCTGTGCC-----CCCTGCTTCCCAAGAGGCGCATTC	1345
QY	1489	TTAAGAAATCTCGACAGCGTGAATCTGTATTACTCTCTCTCCAGAGCCACGAGTCTG	1548
Db	1346	TCAAGAGACCCCGACAGCGCGAGTCTGGCTATCTACTCTCTCCGAGCCAGTGAATCTG	1405
QY	1549	GGGAACTTTTGAAGGCAAGTATGTGTTTTGAATGGGGAACCCCGTGGACAGAAATCTC	1608
Db	1406	GGGAGCTTCTTGAAGCAGCAGCAAGCTGTGTGAAGTGGGAACTCCAGAGGACAGAAAGCTCTC	1465
QY	1609	CACAGGCTTTCAGGGGCTCTCTCCACCGCAAGGGGCAATTCGAAACTCAATGGCAAGTTCT	1668
Db	1466	CGAAGCTTTCAGGGGCTCTCTCTCCATCGCAAGAGCATCTCAAACTCAATGGCAAGTTCT	1525
QY	1669	CCCGCACAGCTTGAAGAGCACTAACCCCTTAAGCACCTTTGGCTCTCCCTGGACCAATGTGCT	1728
Db	1526	CCGACAGAGCTTGAAGAGCTGCGGGCCCAACCACTTGGCTCTCTCTGGATGAATCTCGGCC	1585
QY	1729	CCTCCCATCTGTGACGCCCGCTCCAGCGGCCCTCTCAAGGGGCTGTGATGTAGGACAGCATCC	1788
Db	1586	CACCTCGCCCCCTGAGCCCGGGCCAGGCCACCTCTAGAGGGGCTGTGATGTAGGACAGCATCC	1645
QY	1789	TGTCTCCGAGATCCCTTTGACCAATTTGACCTTGGCCGGAACGTCCTCCGAAACCCCACTGA	1848
Db	1646	TGTCTCTGAGATCCCTTTGACCAAGCTTGAATTTGGCTTGAACGGCTCTCCAGAGCCCCCACTGC	1705
QY	1849	GGGGCTGTGTCTGTGACCAACTGAAGGGGGCTTGAAGCAGCTCTCCCTCAGAG-----	1902
Db	1706	GGGGCTGTGTGTCTGTGACCAACTCAAGGGGCTTGAAGAACCCCTCTCAGAGGGCCCTG	1765
QY	1903	-----GTCTGAAGCATGTGTGGAGGAATCTTTGGGGGAATAGCTGCTTTTCTGTACAG	1956
Db	1766	GAACTCTGTAAGGCGCTGGCGGACAGATCTTTTGGGGGACAGCTGCTTTTCTCTGACAG	1825
QY	1957	ACTGCAAGAGGTATCTGACGCTTACAGACCAAGCCCTAAGAAATGTGCTCAAGCTCAGC	2015
Db	1826	ACTGCAAGAGGTATCAAGCACTTACCGACAGGCACTGAAGGGTCTGTCTCAAGCTCAGC	1884

Search completed: May 16, 2005, 09:29:15
Job time : 1661.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 11:55:01 / Search time 464.804 Seconds
(without alignments)
10216.077 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902

Sequence: 1 cactagtcgattcccaagaat.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
6: /cgn2_6/ptodata/1/ina/6CTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/6CTUS COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	4	US-09-579-664B-4
2	2902	100.0	2902	4	US-10-355-975A-4
3	1372.2	47.3	2501	4	US-09-799-451-215
4	204.2	7.0	2224	4	US-09-949-016-2384
5	204.2	7.0	2224	4	US-09-949-016-2385
6	204.2	7.0	2950	4	US-09-949-016-1546
7	204.2	7.0	2950	4	US-09-949-016-1547
8	202.6	7.0	2175	4	US-09-984-890-1
9	202.6	7.0	2175	4	US-10-274-194-1
10	189.2	6.5	484	4	US-09-270-767-1257
11	189.2	6.5	484	4	US-09-270-767-16539
12	186.4	6.4	2908	3	US-09-930-181-1
13	182.4	6.3	2652	3	US-08-557-006C-39
14	182.4	6.3	2761	3	US-08-557-006C-24
15	181.2	6.2	1742	3	US-08-557-006C-38
16	179.8	6.2	2698	2	US-08-677-298-1
17	179.8	6.2	2912	4	US-09-949-016-2363
18	178.2	6.1	2914	4	US-09-949-016-343
19	176.4	6.1	2899	4	US-09-774-528-112
20	174	6.0	3364	3	US-09-930-181-3
21	173	6.0	2334	4	US-10-003-690-3
22	173	6.0	2337	4	US-10-116-326-1
23	173	6.0	2980	4	US-10-003-690-1
24	172.2	5.9	1539	4	US-09-633-328B-1
25	172	5.9	3609	4	US-09-799-875-6
26	172	5.9	5983	4	US-09-799-875-4
27	171.4	5.9	1647	3	US-09-101-146-44

28	165.6	5.7	1929	3	US-09-359-161-4	Sequence 4, Appli
29	163.6	5.6	1747	3	US-08-357-006C-44	Sequence 44, Appli
30	161.6	5.6	1863	4	US-09-949-016-4303	Sequence 4303, Ap
31	160.4	5.5	1507	3	US-09-523-849-4	Sequence 4, Appli
32	160.4	5.5	1544	3	US-09-734-673-1	Sequence 1, Appli
33	160.4	5.5	2060	3	US-09-523-849-1	Sequence 1, Appli
34	158	5.4	1736	3	US-08-557-006C-37	Sequence 37, Appli
35	158	5.4	1783	3	US-08-557-006C-36	Sequence 36, Appli
36	155.6	5.4	1051	3	US-09-523-849-3	Sequence 5, Appli
37	154	5.3	2112	4	US-10-116-326-5	Sequence 1, Appli
38	149.4	5.1	1599	3	US-09-256-465-1	Sequence 3, Appli
39	149.4	5.1	1599	3	US-09-167-322-3	Sequence 1004, Ap
40	149.4	5.1	1599	4	US-09-023-655-1004	Sequence 3, Appli
41	147.8	5.1	1254	4	US-09-590-740-3	Sequence 70, Appli
42	144.4	5.0	2181	4	US-09-417-197-70	Sequence 138, App
43	144.4	5.0	2184	4	US-09-417-197-138	Sequence 1, Appli
44	144.4	5.0	2610	2	US-09-212-771-1	Sequence 1, Appli
45	144.4	5.0	2610	3	US-09-091-058-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-579-664B-4
Sequence 4, Application US/09579664B
Patent No. 6514719
GENERAL INFORMATION:
APPLICANT: Immunex Corporation
APPLICANT: Bird, Timothy A.
APPLICANT: Martin, G. Duke
APPLICANT: Martin, Unja
APPLICANT: Anderson, Dirk M.
TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
FILE REFERENCE: 2923-A
CURRENT APPLICATION NUMBER: US/09/579, 664B
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2902
TYPE: DNA
ORGANISM: Mus musculus
US-09-579-664B-4

Query Match	100.0%	Score 2902;	DB 4;	Length 2902;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2902;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CAC TAGT GAT CCA A G A T T C G G C A C G A G G C G T G C G G T G G C G T G A C C T T G A G	60	
DB	1	CAC TAGT GAT CCA A G A T T C G G C A C G A G G C G T G C G G T G G C G T G A C C T T G A G	60	
QY	61	CCC GCG GTC A G C G C G G C T G A C T G T G C C G A C C A C T T C A C C T T G C G G T C C G C A	120	
DB	61	CCC GCG GTC A G C G C G G C T G A C T G T G C C G A C C A C T T C A C C T T G C G G T C C G C A	120	
QY	121	CC A G A G T C G T G G C C T T A C T T C A G G C C C G A G C C A G G C T C C T T G G C C T C C G C C T G G	180	
DB	121	CC A G A G T C G T G G C C T T A C T T C A G G C C C G A G C C A G G C T C C T T G G C C T C C G C C T G G	180	
QY	181	C C T G G A G A G G C C C G G C G T G G C G A G G G C C A T C A A G T C G C T T A A C C T T G A T G A	240	
DB	181	C C T G G A G A G G C C C G G C G T G G C G A G G G C C A T C A A G T C G C T T A A C C T T G A T G A	240	
QY	241	A G A A G C A G C G G T A A G G G C A C C A T C A C A A C C T T G C G G C C A G A G T T C C	300	
DB	241	A G A A G C A G C G G T A A G G G C A C C A T C A C A A C C T T G C G G C C A G A G T T C C	300	
QY	301	T G G A G A G C T G G G A A G G G C A C C T A C G G A A G T G A A G A A G G C A G A G A G C T G G G G C	360	
DB	301	T G G A G A G C T G G G A A G G G C A C C T A C G G A A G T G A A G A A G G C A G A G A G C T G G G G C	360	

Dh 2521 CTTAGTACCAAGATGTGAGTGTCTCTGTTTCTAGGCGCAAGGCACTTAAATGTGC 2580
Qy 2581 TGACTTATTTATTTTGTGATTTCTACCTCTGTTTGTGTTTGTGTTTGTGTTG 2640
Dh 2581 TGACTTATTTATTTTGTGATTTCTACCTCTGTTTGTGTTTGTGTTTGTGTTG 2640
Qy 2641 TTTTGTGTTTAAAGTAAATTTTGTCTCTTCAATTAATGTAAATGTGTGTCTGAGAAC 2700
Dh 2641 TTTTGTGTTTAAAGTAAATTTTGTCTCTTCAATTAATGTAAATGTGTGTGTGAGAAC 2700
Qy 2701 TCCATGTGCGACGAGAGTTTATGTATGACAGAGAGATTTGGCAATGATGCTCCCTATTC 2760
Dh 2701 TCCATGTGCGACGAGAGTTTATGTATGACAGAGAGATTTGGCAATGATGCTCCCTATTC 2760
Qy 2761 AAGGAGGAGTGGGAGGCTTTTCAATGATGTCTTGAGCACTGTCTGATTTGAGTCTCA 2820
Dh 2761 AAGGAGGAGTGGGAGGCTTTTCAATGATGTCTTGAGCACTGTCTGATTTGAGTCTCA 2820
Qy 2821 GTCCCTTCAACCAAGGCTGGCCACCTCTCTCATCTTCATCTGTGTGCAAAAAAAA 2880
Dh 2821 GTCCCTTCAACCAAGGCTGGCCACCTCTCTCATCTTCATCTGTGTGCAAAAAAAA 2880
Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
Dh 2881 AAAAAAAAAAAAAAAAAAAAAA 2902

RESULT 2
US-10-355-975A-4
Sequence 4, Application US/10355975A
Patent No. 6759223
GENERAL INFORMATION:
APPLICANT: Immunex Corporation
APPLICANT: Bird, Timothy A.
APPLICANT: Vitco, G. Duke
APPLICANT: Martin, Unja
APPLICANT: Anderson, Dirk M.
TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
FILE REFERENCE: 2923-B
CURRENT APPLICATION NUMBER: US/10/355, 975A
NUMBER OF FILING DATE: 2003-01-30
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2902
TYPE: DNA
ORGANISM: Mus musculus
US-10-355-975A-4

Query Match 100.0%; Score 2902; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACTAGTGAATCCAAAGATTCGGACAGAGGCTGTGGGTGGCGCTGTGACCTTGAG 60
Dh 1 CACTAGTGAATCCAAAGATTCGGACAGAGGCTGTGGGTGGCGCTGTGACCTTGAG 60
Qy 61 CCGCGGCTCAGCGCGGCTCTACTGTGTGCGGACCACTCACTGTGCGGTCCCGCA 120
Dh 61 CCGCGGCTCAGCGCGGCTCTACTGTGTGCGGACCACTCACTGTGCGGTCCCGCA 120
Qy 121 CCATGAGTGGGTGGCTTACTTCCAGGCGCCGAGCGAGGCTCCCTCGGCTCCGCGCTGG 180
Dh 121 CCATGAGTGGGTGGCTTACTTCCAGGCGCCGAGCGAGGCTCCCTCGGCTCCGCGCTGG 180
Qy 181 CTTGAGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTCCCTTAAACCTTGATGA 240
Dh 181 CTTGAGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTCCCTTAAACCTTGATGA 240
Qy 241 AGAAGAGAGGAGTGAAGCGGACATCAAAACAACCTGCGGACCGCTTACGAGTTCC 300
Dh 241 AGAAGAGAGGAGTGAAGCGGACATCAAAACAACCTGCGGACCGCTTACGAGTTCC 300

Qy 301 TGAAGAGCTGGGCGAAGGCGCACTTACGGAAAGTGAAGAGGCAAGAGAGCTCGGGGC 360
Dh 301 TGAAGAGCTGGGCGAAGGCGCACTTACGGAAAGTGAAGAGGCAAGAGAGCTCGGGGC 360
Qy 361 GTCTGGGTGCATCAATGCTTCAAGAGAGCAAAATCAAGAGAGAGATCTGCTGC 420
Dh 361 GTCTGGGTGCATCAATGCTTCAAGAGAGCAAAATCAAGAGAGAGATCTGCTGC 420
Qy 421 ACATAGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCACTATTCATCC 480
Dh 421 ACATAGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCACTATTCATCC 480
Qy 481 ATGAAGTGTGAGATGACAGCAAGATGTGATGTTCATGAGATGACCGGAGCG 540
Dh 481 ATGAAGTGTGAGATGACAGCAAGATGTGATGTTCATGAGATGACCGGAGCG 540
Qy 541 ATCTGTATGATTAATCATGAGAGGCGCAAGGCTGATGAGGCGGAGCGGAGATTTCT 600
Dh 541 ATCTGTATGATTAATCATGAGAGGCGCAAGGCTGATGAGGCGGAGCGGAGATTTCT 600
Qy 601 TCCGACAGATCTGTCTGCGCTGCACTACTGCAACAGAGGAGATGTTCAACGAGATC 660
Dh 601 TCCGACAGATCTGTCTGCGCTGCACTACTGCAACAGAGGAGATGTTCAACGAGATC 660
Qy 661 TCAAGCTGGAAGACATCTCTTATGATGCAATGAGAAATCATCAAGATTTGAGCC 720
Dh 661 TCAAGCTGGAAGACATCTCTTATGATGCAATGAGAAATCATCAAGATTTGAGCC 720
Qy 721 TCTCAACCTGTACCAAAAGGCAAGTCTCTCCAGAGCTTCTGTGGAGCCCTCTTACG 780
Dh 721 TCTCAACCTGTACCAAAAGGCAAGTCTCTCCAGAGCTTCTGTGGAGCCCTCTTACG 780
Qy 781 CCTGCGCTGAGATGATGACAGGAGGCGCTATGAGGCGGAGGAGGAGGAGGAGTCTC 840
Dh 781 CCTGCGCTGAGATGATGACAGGAGGCGCTATGAGGCGGAGGAGGAGGAGGAGTCTC 840
Qy 841 TGGCGGTTCTCTGTACATCTGTGTGATGAGCAACATGCTTTTGAAGGAGAGATCAT 900
Dh 841 TGGCGGTTCTCTGTACATCTGTGTGATGAGCAACATGCTTTTGAAGGAGAGTATA 900
Qy 901 AAACACTGTGAGCAAAATCATGTAACGAGGCTTACCGTGAACCCCAAGCGCTCCGATG 960
Dh 901 AAACACTGTGAGCAAAATCATGTAACGAGGCTTACCGTGAACCCCAAGCGCTCCGATG 960
Qy 961 CCGTGGCTGATCGGAGGCTGTTAATGAGAACCCCAACCCCGGAGGAGCACTGAGAG 1020
Dh 961 CCGTGGCTGATCGGAGGCTGTTAATGAGAACCCCAACCCCGGAGGAGCACTGAGAG 1020
Qy 1021 ATGTAGCAGTCAATGTGGGTCAACTGAGGTTACACCAAGGAGTGGGAGACAGAGAG 1080
Dh 1021 ATGTAGCAGTCAATGTGGGTCAACTGAGGTTACACCAAGGAGTGGGAGACAGAGAG 1080
Qy 1081 CCTGCGTGAAGGTGGACCTTATGTGTGACTTTGACCGGAGCTTCAATGAGGAGTGT 1140
Dh 1081 CCTGCGTGAAGGTGGACCTTATGTGTGACTTTGACCGGAGCTTCAATGAGGAGTGT 1140
Qy 1141 TACGTGCTCTCTCGGCGCCCTCTGTGAGATGAGGCAAGGTGTGAGCTTCTTCAAGC 1200
Dh 1141 TACGTGCTCTCTCGGCGCCCTCTGTGAGATGAGGCAAGGTGTGAGCTTCTTCAAGC 1200
Qy 1201 AGCAGTGGCGGAGTGAAGCACTTACCTGAGGCTGAGCGGACCACTTCTTAAAG 1260
Dh 1201 AGCAGTGGCGGAGTGAAGCACTTACCTGAGGCTGAGCGGACCACTTCTTAAAG 1260
Qy 1261 AGTCCGAAAGAGATGACATGTCTCAAAATCTGCAAGGTGACCGGCTGAGGATCT 1320
Dh 1261 AGTCCGAAAGAGATGACATGTCTCAAAATCTGCAAGGTGACCGGCTGAGGATCT 1320
Qy 1321 CTTCTGCGCTGTGCAAGAGAGCTTAAAGCTTCCGAAGAGATTTCTCAAGAAAGGCT 1380
Dh 1321 CTTCTGCGCTGTGCAAGAGAGCTTAAAGCTTCCGAAGAGATTTCTCAAGAAAGGCT 1380
Qy 1381 CTAAGTGTGAGGAGGATCAGAGAGACCTTCAGAACTCAGACCGGCTGATATCTC 1440

Db 1381 CTACCTGTGACGGGAGGTACAGAGAGACCTCAGAGAACTAGACCCGGTCTGATCTC 1440
Qy 1441 CAGGGCAGCCTGTCCCTGTGTATCCCTGCTCCCAAGGAAAGGATCCTTAAGAAGCTC 1500
Db 1441 CAGGGCAGCCTGTCCCTGTGTATCCCTGCTCCCAAGGAAAGGATCCTTAAGAAGCTC 1500
Qy 1501 GACGCGGTGATCTGTGTACCTCTCTCAGAGGCCAGGAGTCTGGGAACTCTTAG 1560
Db 1501 GACGCGGTGATCTGTGTACCTCTCTCAGAGGCCAGGAGTCTGGGAACTCTTAG 1560
Qy 1561 ACCGCACTGATGTGTGTGTAGTGGGACCCCGTGGAGCAGAGTCTCAAGGCTTCAAG 1620
Db 1561 ACCGCACTGATGTGTGTGTAGTGGGACCCCGTGGAGCAGAGTCTCTCAAGGCTTCAAG 1620
Qy 1621 GGTCTCTCTTCACCGGAGGAGGATCTCAAACTCAATGGCAATTTCTCCGCAAGCT 1680
Db 1621 GGTCTCTCTTCACCGGAGGAGGATCTCAAACTCAATGGCAATTTCTCCGCAAGCT 1680
Qy 1681 TAGAAGGCACTACCCCTTAGCACCTTTGGCTCCCTGGACCACTGGCTCTCCCATCTG 1740
Db 1681 TAGAAGGCACTACCCCTTAGCACCTTTGGCTCCCTGGACCACTGGCTCTCCCATCTG 1740
Qy 1741 CAGCCCGGCGCAGCGCCCTCAGGGGCTGTGTAGTGGAGCAGAGTCTCTCCAGT 1800
Db 1741 CAGCCCGGCGCAGCGCCCTCAGGGGCTGTGTAGTGGAGCAGAGTCTCTCCAGT 1800
Qy 1801 CCTTGAACCAATTGGAATTTGCTGTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
Db 1801 CCTTGAACCAATTGGAATTTGCTGTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGT 1860
Qy 1861 CTGTGACCAACTGTAGGGGGGTTGAGAGAGCTCCCTCAGAGAGGTCTGAAGGATGTGTGC 1920
Db 1861 CTGTGACCAACTGTAGGGGGGTTGAGAGAGCTCCCTCAGAGAGGTCTGAAGGATGTGTGC 1920
Qy 1921 AGGAATCCTTGGGGGATAGCTGCTTTCTCTGACAGACTGTGCAAGGATGTGACAGCT 1980
Db 1921 AGGAATCCTTGGGGGATAGCTGCTTTCTCTGACAGACTGTGCAAGGATGTGACAGCT 1980
Qy 1981 ACAGACAAGCCCTTAGAATCTGCTCAAAAGCTCAGCTGAGAGAGGAGATGTGTCCTAGT 2040
Db 1981 ACAGACAAGCCCTTAGAATCTGCTCAAAAGCTCAGCTGAGAGAGGAGATGTGTCCTAGT 2040
Qy 2041 ATGGGGTATAGCTCTGAGAGGGTTTTCAGAGAGAACCTGGGTGGATTCCTCAAGTAATA 2100
Db 2041 ATGGGGTATAGCTCTGAGAGGGTTTTCAGAGAGAACCTGGGTGGATTCCTCAAGTAATA 2100
Qy 2101 GAGTACATCAAGGGCTCTAGCTGTGACGCTGACCTGAACCTGAAGATGAGAGAAATGCG 2160
Db 2101 GAGTACATCAAGGGCTCTAGCTGTGACGCTGACCTGAACCTGAAGATGAGAGAAATGCG 2160
Qy 2161 ATTGATGTGAGAAAGAAATGGAGACCTTGTGCTGCCGAGTGTATAGTGGGTGCTGAA 2220
Db 2161 ATTGATGTGAGAAAGAAATGGAGACCTTGTGCTGCCGAGTGTATAGTGGGTGCTGAA 2220
Qy 2221 GGTGCTTACCTCTTTTGTGCTATAGTGTCAACCATGATTCCTGCTGTGG 2280
Db 2221 GGTGCTTACCTCTTTTGTGCTATAGTGTCAACCATGATTCCTGCTGTGG 2280
Qy 2281 CTGCACTTCAATAGATTTCTGTTTCATCAACCAAGGGTTAGAACCTGACTTCCCT 2340
Db 2281 CTGCACTTCAATAGATTTCTGTTTCATCAACCAAGGGTTAGAACCTGACTTCCCT 2340
Qy 2341 GGGAGGTATGTGTAGTACTGCTGCAATTTAGAGAGGAAACAGCTCTGTTTCAATCT 2400
Db 2341 GGGAGGTATGTGTAGTACTGCTGCAATTTAGAGAGGAAACAGCTCTGTTTCAATCT 2400
Qy 2401 CTGTGCTGTGATCTCAAAAGACTGTGAGAGACTGCGACCGCTTTGACTTCACTCAA 2460
Db 2401 CTGTGCTGTGATCTCAAAAGACTGTGAGAGACTGCGACCGCTTTGACTTCACTCAA 2460
Qy 2461 GGGGACAGATGCCCTGAGACCCCATTTAGATCTGAGAGCTTGAAGCTGTT 2520

Db 2461 GGGGACAGATGCCCTGAGACCCCATTTAGATCTGAGAGCTTGAACCTTGAAGCTGTT 2520
Qy 2521 CTTAGTACCAGATGTGATGTAGTCTGTGTTTCTCAGGCCAACGGGACCTAGATGTGC 2580
Db 2521 CTTAGTACCAGATGTGATGTAGTCTGTGTTTCTCAGGCCAACGGGACCTAGATGTGC 2580
Qy 2581 TGACTTATTTATTTTGTGATTTCTCACTTCTGTTTGTGTTTGTGTTTGTGTTG 2640
Db 2581 TGACTTATTTATTTTGTGATTTCTCACTTCTGTTTGTGTTTGTGTTTGTGTTG 2640
Qy 2641 TTTTGTTTTAAAGTAAATTTGCTGTTCAATTAATGTAATGCTGTGTTGGGGAAC 2700
Db 2641 TTTTGTTTTAAAGTAAATTTGCTGTTCAATTAATGTAATGCTGTGTTGGGGAAC 2700
Qy 2701 TCCACTGTGCACTGAAGTTATGTACAGAGAAATTTGGAGATGTGCTCTATTC 2760
Db 2701 TCCACTGTGCACTGAAGTTATGTACAGAGAAATTTGGAGATGTGCTCTATTC 2760
Qy 2761 AAGGGGGGTGGGGGCTTTTCAAAATGTATGTGAGCACTGTGGATTGAGTCTCA 2820
Db 2761 AAGGGGGGTGGGGGCTTTTCAAAATGTATGTGAGCACTGTGGATTGAGTCTCA 2820
Qy 2821 GTCCCTTCAACCAAGGCTGGCCACCTCTCATCTTCACTGTGTGCCCCAAAAA 2880
Db 2821 GTCCCTTCAACCAAGGCTGGCCACCTCTCATCTTCACTGTGTGCCCCAAAAA 2880
Qy 2881 AAAAAAAAAAAAAAAAAA 2902
Db 2881 AAAAAAAAAAAAAAAAAA 2902

RESULT 3
US-09-799-451-215
; Sequence 215, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-09-799-451-215

Query Match 47.3%; Score 1372.2; DB 4; Length 2501;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTAAGTCTGCGCCGACCTTCCGAGTCCGCGACCTGAGAGTCCGAGCTTAC 141
DB 65 CTAATGATGTCCTCCGCGCCGCTTGTCTACCTCTGCTGCGCAATGAGTCCGAGTTTTCG 124
QY 142 TCCAGCGCCGAGCGAGAGTCTCCCTGCGCTTCCGCTTGGCTCGGAGAGCGCCGCGCC 201
DB 125 CGCGCGCTCCGCGCCGCTCCCTGCGCGAGAGCTA-----GCCCGCGCG 172
QY 202 TGGCGAGCGGAGTCAATGAGTCCGCTAATCTGATGAGAGAGCAGCGAGTGAAGCGC 261
DB 173 TGGCGAGAGGAGTCAATGAGTCCGCTAATGAGAGAGCAGCGAGTGAAGCGC 232
QY 262 ACCATCAAAACAAACCTGCGGACCGCTAAGAGTCTGAGAGCGCTGCGGCAAGGCA 321
DB 233 ACCACCAAGACAACTGCGGACCGCTAAGAGTCTGAGAGCGCTGCGGCAAGGCA 292
QY 322 CTTACGCGAGAGTGAAGAGCGAGAGAGCTCGGCGCTCTGCTGCGCAATCAAGTCA 381
DB 293 CTTACGCGAGAGTGAAGAGCGAGAGAGCTCGGCGCTCTGCTGCGCAATCAAGTCA 352
QY 382 TCGAGAAAGCAAAATGAAGATGAAGAGATCTGCTGCAATACGAGAGGAGATTGAGA 441
DB 353 TCGAGAAAGCAAAATGAAGATGAAGAGATCTGCTGCAATACGAGAGGAGATTGAGA 412
QY 442 TCAATGCTCACTCAACCAACCCCAATCAATGCTCAATGAAGTCTTGAAGATGAGA 501
DB 413 TCAATGCTCACTCAACCAACCCCAATCAATGCTCAATGAAGTCTTGAAGATGAGA 472
QY 502 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 473 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532
QY 562 AGGCGCGACCGCTGAG 621
DB 533 AGGCGCGACCGCTGAG 592
QY 622 TGCATCTGCTCAACCAACCCCAATCAATGCTCAATGAAGTCTTGAAGATGAGA 681
DB 593 TGCATCTGCTCAACCAACCCCAATCAATGCTCAATGAAGTCTTGAAGATGAGA 652
QY 682 TAGATGCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 741
DB 653 TAGATGCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 712
QY 742 GCAAGTCTCTGAG 801
DB 713 GCAAGTCTCTGAG 772
QY 802 GGAAGCTCTATGAG 861
DB 773 GGAAGCTCTATGAG 832
QY 862 TGGTGCATGAGCAACCATGCTTTGACGAGAGATCAATAAACAATGAGAGAGAGAG 921
DB 833 TGGTGCATGAGCAACCATGCTTTGACGAGAGATCAATAAACAATGAGAGAGAGAG 892
QY 922 GTAAAGGAGCTTACGAG 981
DB 893 GCAAGCGGAGCTTACGAG 952
QY 982 TGTAAATGAG 1041
DB 953 TGTAAATGAG 1012
QY 1042 TCAAATGAG 1101
DB 1013 TCAAATGAG 1072
QY 1102 CTAATGAG 1161
DB 1073 CTGAG 1132

QY 1162 TCTGAGAGATGAG 1221
DB 1133 TCTGAGAGATGAG 1192
QY 1222 GCACTGATCTGAG 1281
DB 1193 GCACTGATCTGAG 1252
QY 1282 TGGCTCAAAATCTGAG 1341
DB 1253 TGGCTCAAAATCTGAG 1312
QY 1342 GCTTAAAGCTTCCGAG 1401
DB 1313 ACCTCAAGCTGAG 1372
QY 1402 AG 1461
DB 1373 AG 1426
QY 1462 TATCCCTGCTCCGAG 1521
DB 1427 ---CCTGCTCCGAG 1483
QY 1522 ACTCTCTCCGAG 1581
DB 1484 ACTCTCTCCGAG 1543
QY 1582 GTGGGAG 1641
DB 1544 GTGGGAG 1603
QY 1642 GCAATCTCAAAATCTGAG 1701
DB 1604 GCAATCTCAAAATCTGAG 1663
QY 1702 CTTTGGCTCCGAG 1761
DB 1664 CTTTGGCTCCGAG 1723
QY 1762 CAGGAGCTGAG 1821
DB 1724 CAGGAGCTGAG 1783
QY 1822 CTGAGAGCTTCCGAG 1881
DB 1784 CTGAGAGCTTCCGAG 1843
QY 1882 TTGAGCAGCTCCCTCAG 1929
DB 1844 TTGAGCAGCTCCCTCAG 1903
QY 1930 TGGGAGATGAG 1989
DB 1904 TGGGAGATGAG 1963
QY 1990 CCTTGAAGATCTGCTCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2049
DB 1964 CACTGAGAGATCTGCTCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2022
QY 2050 GCTCTGAG 2090
DB 2023 GCTCTGAG 2063

RESULT 4
US-09-949-016-2384
; Sequence 2384, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2384
LENGTH: 2224
TYPE: DNA
ORGANISM: Human
US-09-949-016-2384

Query Match 7.0%; Score 204.2; DB 4; Length 2224;
Best Local Similarity 55.0%; Pred. No. 3.6e-42;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 290 CTACGAGTCTCTGGAGACGCTGGCGAAGGCACTACGGAAGGTGAAGAGCAGAGA 349
DB 57 CTACCGGCTCTCTCAAGACCATTTGGCAAGGGTAATTTTCCAAAGGTGAAGTGGCCGACACA 116
QY 350 GAGCTCGGGCGTCTGGTGGCCATCAAGTCCATCAGAGAAACAAATCAAGATGAGCA 409
DB 117 CATCTCACTGGGAAAGAGTAGCTGTGAAGATCATTTGACAAAGCTCAACTGAATCTCTC 176
QY 410 GGATCTGCTGCATACGAGAGGAGATTGAGATCATGCTTCACTCAACACCCCCACAT 469
DB 177 CAGCTCCAGAACTATTTCCGCGAAGTAAGAAATTAAGAGTTTGAATCATCTCCACAT 236
QY 470 CATTGCATCATGAAGTGTGTAAGATGACAGAGAAATGATGTGATGAGATGAC 529
DB 237 AGTTAAATTTATTTGAAGATGAGATGAGAGAAACGCTCACTGTCATGAGATGAC 296
QY 530 CAGCGGAGGAGATCTGATGATTAATCATGAGACCGGCAAGCTGATGAGCGGAGCG 589
DB 297 TAGGGGAGAGGATATTTGATTAATCATGAGGCTCATGAGGAGTAAGAAAGAGGCG 356
QY 590 CAGGCAATTTCTCCGAGAGATCGTGTGCTGCTGCACTACGACCAAGAGGGATCGT 649
DB 357 TCGAGCCAAATTCGCGCAGATAGTGTCTGTGTGAGTACGATCAACGAAAGTTATTTGT 416
QY 650 TCACCGAGATCTCAAGCTGAGAAACATCTTTCTAGATGCCAATGAGAAACATCAAGATTGC 709
DB 417 CCAATAGAGACTTAAAGGCAAGAAACCTGCTCTTGAGTGTGATGAAATCAAGATTGC 476
QY 710 TGAATTTGGGCTCTCCAACTGTACACAAAGGCAAGTTCTCCAGAGTTCTGTGGAG 769
DB 477 AGACTTTGGGCTTCAGCAATGATTAATCACTTTGGGAAACAAAGCTGAGCACTTCTGTGGCAG 536
QY 770 CCCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCTATGAGGAGCCGAGAGTGA 829
DB 537 TCCGCTTATGCTGCTCCCAAGACTTTCAGAGGCAAAATTAATGATGAGCCGAGGTGA 596
QY 830 CAGCTGTCTCTGAGGCTTCTCTGTATCATCTGTGTGATGAGCAACATGACCTTTTGA 889
DB 597 TGTGTGAGGCTTGAAGATTAATCTCTATACCTGTGAGGAGATCCCTGCTTTTGAATGG 656
QY 890 GCGAGATCATTAACATCTGTGAGCAATCATGTAAGGAGGCTTAACGCTGAGCCGCCAA 949
DB 657 ACAGAACTTCAAGAGAGCTGCGGAGAAAGGATCTGAGGAGGAAATACCGTATTTCAATCTA 716
QY 950 GCGGTGCGATGCTGTGTG--GCTGATCCGGTGGCTGTTATGAGGAAACCCCAACCGTGG 1006
DB 717 CATGTCCAGGAGCTGTGAAACCTGTCTTAAGAAATTTCTCATTTCTTAATCCCAAGAG 776
QY 1007 GCGCCACATGAGAGATGAGCACTCATTTGTGTGCTCAACTGGGATTAC 1055
DB 777 AGGCACTTTAGAGCAATATGAAAGATGATGATGATGTGGTAC 825

RESULT 5
US-09-949-016-2385
Sequence 2385, Application US/09949016
Patent No. 681239
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2385
LENGTH: 2224
TYPE: DNA
ORGANISM: Human
US-09-949-016-2385

Query Match 7.0%; Score 204.2; DB 4; Length 2224;
Best Local Similarity 55.0%; Pred. No. 3.6e-42;
Matches 423; Conservative 0; Mismatches 343; Indels 3; Gaps 1;

QY 290 CTACGAGTCTCTGGAGACGCTGGCGAAGGCACTACGGAAGGTGAAGAGCAGAGA 349
DB 57 CTACCGGCTCTCTCAAGACCATTTGGCAAGGGTAATTTTCCAAAGGTGAAGTGGCCGACACA 116
QY 350 GAGCTCGGGCGTCTGGTGGCCATCAAGTCCATCAGAGAAACAAATCAAGATGAGCA 409
DB 117 CATCTCACTGGGAAAGAGTAGCTGTGAAGATCATTTGACAAAGCTCAACTGAATCTCTC 176
QY 410 GGATCTGCTGCATACGAGAGGAGATTGAGATCATGCTTCACTCAACACCCCCACAT 469
DB 177 CAGCTCCAGAACTATTTCCGCGAAGTAAGAAATTAAGAGTTTGAATCATCTCCACAT 236
QY 470 CATTGCATCATGAAGTGTGTAAGATGACAGAGAAATGATGTGATGAGATGAC 529
DB 237 AGTTAAATTTATTTGAAGATGAGATGAGAGAAACGCTTACCTGTGATGAGATGAC 296
QY 530 CAGCGGAGGAGATCTGATGATTAATCATGAGACCGGCAAGCTGATGAGCGGAGCG 589
DB 297 TAGGGGAGAGGATATTTGATTAATCATGAGGCTCATGAGGAGTAAGAAAGAGGCG 356
QY 590 CAGGCAATTTCTCCGAGAGATCGTGTGCTGCTGCACTACGCAACCAAGAGGGATCGT 649
DB 357 TCGAGCCAAATTCGCGCAGATAGTGTCTGTGTGAGTACGATCAACGAAAGTTATTTGT 416
QY 650 TCACCGAGATCTCAAGCTGAGAAACATCTTTCTAGATGCCAATGAGAAACATCAAGATTGC 709
DB 417 CCAATAGAGACTTAAAGGCAAGAAACCTGCTCTTGAGTGTGATGAAATCAAGATTGC 476
QY 710 TGAATTTGGGCTCTCCAACTGTACACAAAGGCAAGTTCTCCAGAGTTCTGTGGAG 769
DB 477 AGACTTTGGGCTTCAGCAATGATTAATCACTTTGGGAAACAAAGCTGAGCACTTCTGTGGCAG 536
QY 770 CCCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCTATGAGGAGCCGAGAGTGA 829
DB 537 TCCGCTTATGCTGCTCCCAAGACTTTCAGAGGCAAAATTAATGATGAGCCGAGGTGA 596
QY 830 CAGCTGTCTCTGAGGCTTCTCTGTATCATCTGTGTGATGAGCAACATGACCTTTTGA 889
DB 597 TGTGTGAGGCTTGAAGATTAATCTCTATACCTGTGAGGAGATCCCTGCTTTTGAATGG 656
QY 890 GCGAGATCATTAACATCTGTGAGCAATCATGTAAGGAGGCTTAACGCTGAGCCGCCAA 949
DB 657 ACAGAACTTCAAGAGAGCTGCGGAGAAAGGATCTGAGGAGGAAATACCGTATTTCAATCTA 716


```

Db      824  CCAATGAGACTTAAAGGCAAGAAACCTGCTCTTGATGCTGATATGAAACATCAAGATTGC 883
QY      710  TGAATTGGGCTCTTCAACCTGTAACAAGCAAGTTCCTCAAGAGCTTCTGTGAG 769
Db      884  AGACTTTGGCTTCAGCAATGAAATCACCTTTGGGAAACAGCTGACACCTTCTGTGAG 943
QY      770  CCCTCTAGGCGCTCGGCTGAGATGCAACGGGAAGCCCTATGTGGCCCAAGAGTGA 829
Db      944  TCCCTCTTATGCTGCCCCAGAACTCTTCAGGGCAAAAAATATATGAGACCCGAGTGA 1003
QY      830  CAGCTGTCTCTGAGCGCTTCTCTGTACATCTGTGATGACACCAATGACCTTTGACG 889
Db      1004  TGTGTGAGGCTTGAAGATTATCCCTTATACCTGTGAGCGGATCCCTGCTTTGATGG 1063
QY      890  GCAAGATCATTAACACTGTGTGAAGCAATCATGACGGGGCTTACCGTGAAGCCGCA 949
Db      1064  ACAGAACTCAAGAGCTGCGGGGAAACGGGTACTGAGGGGAAAAATACCGTATTCATTCTA 1123
QY      950  GCGGTCCGATGCGCTGTG---GCCGTATCCGCTGCTGTAAATGTTGAACCCGACCGTGC 1006
Db      1124  CATGTCCACGACTGTGAAAACTGTTAAGAAATTTCTCATTTCTTAATCCCAAGAG 1183
QY      1007  GGCCACACTGAGAGATGTAGCCAGTCAATGTGTGGGTCAACTGGGATTAC 1055
Db      1184  AGGACATTTAGAACAATCATGAAAGATCATGATGATATGTGGTAC 1232

```

```

RESULT 8
US-09-984-890-1
; Sequence 1, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-890-1

```

```

Query Match      7.0%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 9.3e-42;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY      290  CTACGAGTTCTCTGAGACGCTGGCAAGGCACTTACGAGAAAGTGAAGAGCAGAGA 349
Db      156  CTACCGGCTCTCAAGACATTTGGCAAGGTAATTTTCCAGGTGAAGTTGGCCGACA 215
QY      350  GAGCTGGGGCGTCTGTGGCCATCAAGTCCATGAGAAAGCAAAATCAAAGTGAACA 409
Db      216  CATCTGACTGAGGAAAGAGGTAGCTGTGAAGATCATTTGACAAAGTCACTGACTCTC 275
QY      410  GGATCTGCTGCACATAGAGGAGGAGATTGAGATCATTTCACTCAACACCCCAAT 469
Db      276  CAGCTTCAGAAACTATTCGCGAGAGTAAGATATATAGAGTTTGAATCATCCCAAT 335
QY      470  CATTGCCATCATGAAGTGTGTTGAGATAGCAGCAAGATGTGATGTCTCATGAGATGC 529
Db      336  AGTTAATTTATTTGAAGTATGAGACTGAGAAAGCCTCTAAGTTTCAATGAGATGCG 395
QY      530  CAGCCGAGGCGATCTGTATATTAATCATCACTGAGCGGCAACGCTGATGAGCGGAGCG 589
Db      396  TAGTGGGAGAGAGTATTTGATTAATCAATGAGGATGAGAAAGAAAGAGGC 455
QY      590  CAGGCAATTTCTCGAAGATCGTGTCTGCCCTGCACTACTGCAACCAAGAGGATGCT 649

```

```

Db      456  TCGAGCCAATTTCCGCAAGTATGTCTGTGTGCAATCTGTGACCAAGAGTTATTTG 515
QY      650  TCACCGAGATCTCAAGCTGAAAAACATCTTTAGATGCCAATGAAACATCAAGATTGC 709
Db      516  CCAATGAGACTTAAAGCAGAAACCTGCTCTTGATGCTGATATGAAATCAATCAAGATTGC 575
QY      710  TGAATTGGGCTCTTCAACCTGTACCAAGGCAAGCTTCTCAAGAGCTTCTGTGAGAG 769
Db      576  AGACTTTGGCTTCAAGCAATGAATTAACCTTTGGGAAACAGCTGACACCTTCTGTGAG 625
QY      770  CCCTCTAGGCGCTCGCTGAGATGATGCAAGGGAAGCCCTATGTGGGCCCAAGAGTGA 829
Db      636  TCCCTCTTATGCTGCCCCAGAACTTTCAAGGCAAAAAATATATGATGACCCAGAGTGA 695
QY      830  CAGCTGTCTCTGAGCGCTTCTCTGTACATCTGTGATGACACCAATGACCTTTGACG 889
Db      696  TGTGTGAGGCTTGAAGATTATCCCTTATACCTGTGAGCGGATCCCTGCTTTGATGG 755
QY      890  GCAAGATCATTAACACTGTGTGAAGCAATCATGAAACGGGGCTTACCGTGAAGCCGCA 949
Db      756  ACAGAACTCAAGAGCTGCGGGGAAACGGGTACTGAGGGGAAAAATACCGTATTCATTCTA 815
QY      950  GCGGTCCGATGCGCTGTG---GCCGTATCCGCTGCTGTAAATGTTGAACCCGACCGTGC 1006
Db      816  CATGTCCACGACTGTGAAAACTGTTAAGAAATTTCTCATTTCTTAATCCCAAGAG 875
QY      1007  GGCCACACTGAGAGATGTAGCCAGTCAATGTGTGGGTCAACTGGGATTAC 1055
Db      876  AGGACATTTAGAACAATCATGAAAGATCATGATGATATGTGGTAC 924

```

```

RESULT 9
US-10-274-194-1
; Sequence 1, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-274-194-1

```

```

Query Match      7.0%; Score 202.6; DB 4; Length 2175;
Best Local Similarity 54.9%; Pred. No. 9.3e-42;
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY      290  CTACGAGTTCTCTGAGACGCTGGCAAGGCACTTACGAGAAAGTGAAGAGCAGAGA 349
Db      156  CTACCGGCTCTCAAGACATTTGGCAAGGTAATTTTCCAGGTGAAGTTGGCCGACA 215
QY      350  GAGCTGGGGCGTCTGTGGCCATCAAGTCCATGAGAAAGCAAAATCAAAGTGAACA 409
Db      216  CATCTGACTGAGGAAAGAGGTAGCTGTGAAGATCATTTGACAAAGTCACTGACTCTC 275
QY      410  GGATCTGCTGCACATAGAGGAGGAGATTGAGATCATTTCACTCAACACCCCAAT 469
Db      276  CAGCTTCAGAAACTATTCGCGAGAGTAAGATATATAGAGTTTGAATCATCCCAAT 335
QY      470  CATTGCCATCATGAAGTGTGTTGAGATAGCAGCAAGATGTGATGTCTCATGAGATGC 529
Db      336  AGTTAATTTATTTGAAGTATGAGACTGAGAAAGCCTCTAAGTTTCAATGAGATGCG 395
QY      530  CAGCCGAGGCGATCTGTATATTAATCATCACTGAGCGGCAACGCTGATGAGCGGAGCG 589

```

Db 396 TAGTGGCGAGAGGATTTTAACTTAGTGGCTTCATGCGAGATGAAAGAAAAGAGGC 455
Qy 590 CAGGCAATTTCTTCGACAGATCGGTCTGCCCTGCACTACGACCAAGAGGATGCT 649
Db 456 TCGAGCCAAATTCGCCAGAGTATGTCTGCTGACATGCTGACCAAGAGTTATGCT 515
Qy 650 TCACCGAGATCTCAGGTGAGAAACATCTTATGATGCGAATGAGAAACATCAAGATTGC 709
Db 516 CCAATAGAGACTTAAAGGCAAGAAACGCTCTTGATGCTGATATGAAACATCAAGATTGC 575
Qy 710 TGAATTTGGCTCTCCCACTGTAACCAAGGCAAGTTCTCCAGACGTTCTGTGGAG 769
Db 576 AGACTTTGGCTTACGACATGAAATTCACCTTTGGAGAACAGCTGACACCTTGTGGCAG 635
Qy 770 CCCCTCTACGCTCGGCTGATAGTCAACGAGAGCCCTATGTTGGGCCAGAGTTGGA 829
Db 636 TCCCCCTTATGCTGCCCAAGACTCTTCAGGGCAAAAATATATGATGACCCGAGGTGGA 695
Qy 830 CAGCTGCTCTGCGGCGTCTCTCTGTAATCTCTGTCATGTCATGTCACATGCTTTGACG 889
Db 696 TGTGTGAGGCTAGGAGATTAATCTCTATACATGCTGTCAGGGGATCCCTGCTTTGATGG 755
Qy 890 GCAAGATCATPAAACACTGCTGTAAGCAATCATGTAACGGGCTTACCGTGAAGCCGCA 949
Db 756 ACAGAACTCAAGAGAGCTGCGGGAACGGGTACTGAGGGGAAAATACCGTATTCATTCTA 815
Qy 950 GCGGTCCGATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 816 CATGTCCACGACCTGTAAACCTGCTTAAGAAATTTCTCTTATATCCCAAGAGAG 875
Qy 1007 GGCCACACTGAGAGATGATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
Db 876 AGGCACTTGAAGCAATCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 924

RESULT 10
US-09-270-767-1257/c
; Sequence 1257, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1257
; LENGTH: 484
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-1257

Query Match 6.5%; Score 189.2; DB 4; Length 484;
Best Local Similarity 64.3%; Pred. No. 1.2e-38;
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 355 CGGGGCGTCTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAAGATGAGCAGATC 414
Db 442 CCGGCCAGAGAGTGGCTATCAAAACCATCAAGAGTCAAGATCGAGGCCGAGCGGATT 383
Qy 415 TGTGTCACATACGAGGAGATTTGATCTTCACTCAACACCCCACTCATTTG 474
Db 382 TGTGTCGCTATCCGTCGAGAGTGAATTAAGCTCAGTCAATCCCAACATCATCC 323
Qy 475 CCAATCCATGAAGTTTGAAGATAGCAGAGATTTGATGATGATGATGATGATGATGATG 534
Db 322 ACATCTAGAGAGTTTGAAGATGATGAGAAATGCTGCTAGTATGATGATGATGATGATG 263
Qy 535 GAGCGCATCTGTATGATTAATCATCAAGTACCGGCAACGCTGATGATGATGATGATGATG 594
Db 262 GCGCGAGAGCTTCAAGTATCTGTCTGAAGGAAAGTTTCAACGAGAGAGGCGAGAC 203

Qy 595 ATTCTTCCGACAGATGCTGTCTGCTCTGCACTATCTGCAACAGAGAGGATGCTTACC 654
Db 202 GCATCTTCCGCGCAGAGTGGCCACCGCGCTTACTATCTGTCACAGCAAGATCTGCCATC 143
Qy 655 GAAATCTCAAGCTGGAAGAAACATCTTCTATATGCAATGAGAAACATCAAGATTGCTGACT 714
Db 142 GCATCTCAAGCTGGAAGAAACATCTTCTGACAGAGAGGCAATGCTAAGATTGCTGATTT 83
Qy 715 TTGGCTCTCCAACTGTACCAAGAGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTC 774
Db 82 TTGGGTTGTGAATGTCTTGTATGATGACCAAGGACTGCTGGGCACTTTTGGGGTCCCCAC 23
Qy 775 TCTAAGCTCTGCTGAGATAGT 796
Db 22 TCTATGCTCTGCGGAAATTTGT 1

RESULT 11
US-09-270-767-16539/c
; Sequence 16539, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16539
; LENGTH: 484
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-16539

Query Match 6.5%; Score 189.2; DB 4; Length 484;
Best Local Similarity 64.3%; Pred. No. 1.2e-38;
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 355 CGGGGCGTCTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAAGATGAGCAGATC 414
Db 442 CCGGCCAGAGAGTGGCTATCAAAACCATCAAGAGTCAAGATCGAGGCCGAGCGGATT 383
Qy 415 TGTGTCACATACGAGGAGATTTGATCTTCACTCAACACCCCACTCATTTG 474
Db 382 TGTGTCGCTATCCGTCGAGAGTGAATTAAGCTCAGTCAATCCCAACATCATCC 323
Qy 475 CCAATCCATGAAGTTTGAAGATAGCAGAGATTTGATGATGATGATGATGATGATGATG 534
Db 322 ACATCTAGAGAGTTTGAAGATGATGAGAAATGCTGCTAGTATGATGATGATGATGATG 263
Qy 535 GAGCGCATCTGTATGATTAATCATCAAGTACCGGCAACGCTGATGATGATGATGATGATG 594
Db 262 GCGCGAGAGCTTCAAGTATCTGTCTGAAGGAAAGTTTCAACGAGAGAGGCGAGAC 203
Qy 595 ATTCTTCCGACAGATGCTGTCTGCTCTGCACTATCTGCAACGAGAGAGGATGCTGACC 654
Db 202 GCATCTTCCGCGCAGAGTGGCCACCGCGCTTACTACTATCTGTCACAGCAAGATGTCGCAATC 143
Qy 655 GAAATCTCAAGCTGGAAGAAACATCTTCTATATGCAATGAGAAACATCAAGATTGCTGACT 714
Db 142 GCATCTCAAGCTGGAAGAAACATCTTCTGACAGAGAGGCAATGCTAAGATTGCTGATTT 83
Qy 715 TTGGCTCTCCAACTGTACCAAGAGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTC 774
Db 82 TTGGGTTGTGAATGTCTTGTATGATGACCAAGGACTGCTGGGCACTTTTGGGGTCCCCAC 23
Qy 775 TCTAAGCTCTGCTGAGATAGT 796
Db 22 TCTATGCTCTGCGGAAATTTGT 1

RESULT 12
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930.181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match 6.4%; Score 186.4; DB 3; Length 2908;
Best Local Similarity 53.5%; Pred. No. 1.7e-37;
Matches 413; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

```
QY 288 CGCTACGAGTTCTGAGAGCCGTGGCAAGGCGACCTACGGGAAGTGAGAGGACACGA 347
DB 157 CCCTACCGGCTGAGAGACGCTGGCAAGGGGACGACAGGTCTGTGAGAGCTGGGGGTT 216
QY 348 GAGAGCTCGGGGCGTCTGTGGTGGCCATCAATCCATCAGAAAGCAAAATCAAGATGAG 407
DB 217 CACGCGCTCACCTGCCAGAGGTGGCCATCAAGATCTGCAACCGTAGAGAGCTCAGGAG 276
QY 408 CAGAGTCTGCTGCACATACGAGGAGAGATTGATGATCTGTCTTCACTCAACACCCGAC 467
DB 277 TCGGTGCTGATGAGAGTGAGCGGAGATCGCATCTGGAAGCTCATTTAGACACCCGAC 336
QY 468 ATCATTTGCCATCCATGAGATGTTTGAAGATGACAGCAAGATTGATGTCATGAGATAT 527
DB 337 GTCTTAAGCTGACGACGATTTATGAAACAAATAATTTGATCTGTGCTTGAACAC 396
QY 528 GCCAGCGGAGCGATCTGTATGATTAATCATCAGTACGAGCGGACGAGTGAAGCGGAGAC 587
DB 397 GTGTCAGGTGTGAGCTCTTCTGACTGCTGTGAAGAGGAGGCTGACCGCTTAAGAG 456
QY 588 GCCAGGATTTCTTCCGACAGATGTGTGCTGCTGCACTACTCCACACAGAGCGGATC 647
DB 457 GCTCGAAGTTCTTCCGGCAGATCATCTGTGCTGCACTTCTGCAAGCCTCCATA 516
QY 648 GTTACCGAGATCTCAAGCTGGAAAAATCTTTTGAATGCCAATGGAACATCAAGATT 707
DB 517 TGCACAGGAGATCGAAACCTGAAACCTCTGTGACGAGAAACAACATCCGATC 576
QY 708 GCTGACTTTGGCCCTCTCAACCTGTACCAAAAAGCAAGTTCTCCGAGCGTTGTGGG 767
DB 577 GCAGACTTTGGCATGGGTCTCTCAGGTTGGCAGACGCTGTGGAGACGAGCTGGGG 636
QY 768 AGCCCTCTCTACGCTGCTGAGATAGTCAACGAGGAGCCCTTATGTGGGCCAGAGTG 827
DB 637 TCCGCCCACTACGCTGCCCCGAGGTATCCGGGGGAGAAATGACGGCGGAAAGCG 696
QY 828 GACAGCTGTCTCTGGGCGTTCTCTGTATCATCTGTGTGATGGCACCATGCTTTGAC 887
DB 697 GACCTGTGTGAGCTGCGGCGTCACTCTGTCCTGTGCTGGGTGGGGCTGTGCCCTTGAC 756
QY 888 GGGGAGGATCATAAAACTGTGTGAGAGCAATCAAGTAACGGGGGTTCACCGTAGACCCGCC 947
DB 757 GATGACAACTTTCGACAGCTGCTGAGAGAGTGAAGCGGGGCGGTTCACATCCGAC 816
QY 948 ---AAGCGTCCGATGCTGTGGCTGATCCGATGCTGTGAATGATGTAACCCGCGT 1004
DB 817 TTTATCCCGCCGACCTGCAAGAGTCTGTGACGGGGCATGATGAGGTGAGCGCGCACGC 876
```

RESULT 13
US-08-557-006C-39
; Sequence 39, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Bert, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 2652
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
US-08-557-006C-39

Query Match 6.3%; Score 182.4; DB 3; Length 2652;
Best Local Similarity 53.0%; Pred. No. 1.7e-36;
Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;

```
QY 242 GAGACGGCGGTGAAGCGGACATCAACAAACCTGCGGACCGCTACGAGTTCT 301
DB 4 GAACATGCTGAGAGAGACAGACGAGCGGCTGTGAATCGACATCACTGTGCGG 63
QY 302 GAGACCGCTGGGACAGGCGCTACCGGAGAGTGAAGAGCGAGAGCTCG--GG 358
DB 64 GACACCCCTGGGCTGGGACCTTCGCGAAGAGTGAAGATTTGAGAACATTAACAGG 123
QY 359 GCGTCTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAAGATGACGAGATCTGT 418
DB 124 CCATTAAGGTGAGTTAAGATCTTAATATAGACAGAAATTCGAGTTAGATGTTGG 183
QY 419 GCACATACGAGAGGAGATTAAGATCATGTCTTCACTCAACACCCCAATCATTTGCAT 478
DB 184 AAAAATTAACGAGAAATTCATAAATCTTAACTTTTGTGATCTCATATTATCAAACT 243
QY 479 CCATGAAGTTTGAAGATAGCAAGATTGTGATTTGATGTCATGAGATGTCAGCGGAG 538
DB 244 CTACCAAGTATACGACTCCAAACAGACTTTTATATGATATGATATGTGTGAGG 303
QY 539 CGATCTGTATGATTACATCAGTACGAGCGGACGCGCTGAGTGAACGGGACCGCAGGACTTT 598
DB 304 TGAATTGTGACTACATCTGTAAACGCGGAGGTTGAAGAGTGAAGCTGCGCGGCT 363
QY 599 CTTCGACAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
DB 364 CTTCGACAGATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
QY 659 TCTCAAGCTGGAAGAACATCTTCTAGATGTCATGAGTGAAGACATCAAGTATGCTGCTTGG 718
DB 424 CTGAAAGCAGAGAGCTGTGCTGAGCGCCAGATGAATGCTAAGATTAAGCTTCCG 483
```


Db	3	GGCTGAGAAAGCAGAAAGCACGACGGGCGTGTGAAGATCCGACACTACGTGTCTGGGGAGACAC	62	
Oy	308	GCTGGGCAAGGGCACTTACGGGAAAGTGAAGAAAGGCACGAGAGACT---	CGGGCGCT	364
Db	63	CCTGGGCGTGGGCACTTCCGGCAAGTGAAGTTGGAGACATCAATTGACAGGGCCATTAA	122	
Oy	365	GGTGGCCATCAAGTCCATCAGGAAAGACAAAATCAAAGATGACGAGATCTGTCCACAT	424	
Db	123	AGTGGCAATTGAAGATCTTAAATGACAGAAAGTTGGCAGTTTAAATGTTGTGGAAAAAT	182	
Oy	425	ACGGAGGAGATTGATCATGTCTTCACTCAACACCCCAATCATTTGCCATTCATGA	484	
Db	183	AAAACGAAATTTCAAATCTTTAACTCTTTCGTCACTCTCAATTAATCAAACTTTACA	242	
Oy	485	AGTGTTTGAGAAATGACAGCAAGATTGTGATTTGTCAATGAGATATGCCAGCGAGCATCT	544	
Db	243	AGTATTCAGCACTCCACAGACTTTTATGTGTAATGGAATATGTGTCTGGAGGGTAAT	302	
Oy	545	GTATGATTTACATCACTGAGCGGGCCACGGCTGATGAGCGGGAGCGCAGGCAATTTCTTCG	604	
Db	303	GTTCGACTTACATCTGTAAACACGGGAGGGTTGAAGAGGTGGAAAGCTCGCCGGCTTCCA	362	
Oy	605	AACAGTGGTGTCTGCCCTGCACTAATGGCAACGAAACGGATTCGTTACCGAATCTCA	664	
Db	363	GCAAGTCTGTCTGCGGTGAGCTACTGTTCACAGGCAATGGTTGTCCACAGGAACTTAA	422	
Oy	665	GCTGAAAAAATCCCTTCTAGATATGCCAATGSAACATCAAGATTGCTGACTTTGGCCTTC	724	
Db	423	GCCAGAGAACGTGTGGTGTGAGCGCCCAATGAATCTTAAGTACTGTGACTTCGACCTTC	482	
Oy	725	CAACCTGTACCAAAAGGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTTACGCTC	784	
Db	483	TAAATATGATGCAGATGGTGAATTTCTACGAATCTAGCTGTGGATGCCAAATTAATGAGC	542	
Oy	785	GCTGAGATATGTCACACGGGAAGCCCTAATGTGGGCCCAAGGTGACAGCTGTCTCTGGG	844	
Db	543	ACCGAGGTCATCTTCAGAGAGCGTGAATGCGGCTCTAAGGTGATTAATCTGAGCTGTGG	602	
Oy	845	CGTTCTCTGTACATCCTGTGTCATGAGCAACATGCGCTTTGACGGGACGAGATCATTAAC	904	
Db	603	TGTTATCTCTGTATGCGCTTCTTCTGTGGACCTCCCGTTTCGACGATGACAGTGGCTTAC	662	
Oy	905	ACTGTGAAGCAAAATCACTTAAACGGGGCTTACCGGTAGCGCCCAAGCCGTCCGATGCTG	964	
Db	663	GCTCTTTAAGAAAGATCCGAGGGGGTGTGTCTTACATCCCGGAGTATCTCAACCGTTTAT	722	
Oy	965	TGGC---CTGATCCGGTGGCTGTTAATGTGTGAACCCACCCGTGGGCGACACTGAGGA	1021	
Db	723	TGCCACTCTGCTGATGCAATGCTGCAAGTGGACCCCTTTGAAGCGAGCAATCATCAAGA	782	
Oy	1022	TGTAGCAGCATTTGTGGGTCACTGGGGTTACCA	1059	
Db	783	CATACGAGCAATGAATGTTTAAACAGAAATTTGCCA	820	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2005, 22:57:45 ; Search time 1510.61 Seconds
(without alignments)
11372.259 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902
Sequence: 1 cactagtgcattcccaagat.....aaaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_16Dec04:*
2: geneseq1980s:*
3: geneseq1990s:*
4: geneseq2000s:*
5: geneseq2001as:*
6: geneseq2001bs:*
7: geneseq2002as:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	4 AAC90433	AAC90433 Murine Ly
2	2863.4	98.7	3073	13 ABD33083	ABD33083 Murine ca
3	2066.8	71.2	2929	6 ABD31710	Aad31710 Rat SNF1/
4	1938.8	66.8	37278	13 ABD33082	ABD33082 Murine ca
5	1379	47.5	3353	6 ABK14000	ABK14000 CDNA enco
6	1375.8	47.4	3395	12 ADL14160	ADL14160 Novel hum
7	1375.8	47.4	3395	12 AAK94280	AAK94280 Human ful
8	1375.8	47.4	3395	12 ADL30885	ADL30885 Full leng
9	1375.2	47.4	3404	13 ABD33085	ABD33085 Human can
10	1374.2	47.4	3443	5 ABX71420	ABX71420 Human cel
11	1374.2	47.4	3443	10 ADF76964	ADF76964 Novel hum
12	1374.2	47.4	3443	10 ADF81952	ADF81952 Leukaemia
13	1374.2	47.4	3443	12 ADO20171	ADO20171 Human PRO
14	1374.2	47.4	3463	12 ADJ96554	ADJ96554 Human cal
15	1372.2	47.3	2501	12 ADM43851	ADM43851 Novel hum
16	1371.8	47.2	2043	6 ABZ11333	ABZ11333 Human pol
17	1370.6	47.2	3360	6 AAD26459	AAZ26459 Human kin
18	1360.2	46.9	1884	12 ADL14162	ADL14162 Novel hum
19	1349.2	46.5	2291	4 AAF44659	AAf44659 Novel pro
20	1349.2	46.5	2291	12 ADI29357	ADI29357 Human MAR

21	1323.2	45.6	3200	4 AAF75338	AAf75338 Human TGF
22	1230	42.4	1833	10 ABZ77163	ABz77163 Human pro
23	849.2	29.3	2616	6 ABQ72599	ABq72599 Human MDD
24	849.2	29.3	2619	6 ABQ72698	ABq72698 Human MDD
25	732	25.2	39699	-13 ABD33084	ABD33084 Human can
26	507.6	17.5	6828	4 AAD14327	AAZ14327 Human pro
27	507.6	17.5	6828	10 ADE38420	ADe38420 Human pro
28	507.6	17.5	6828	11 ADN95767	ADn95767 Human BPC
29	507.6	17.5	6828	12 ADL74808	ADl74808 Marker ge
30	507.6	17.5	6828	12 ADL25352	ADl25352 Human ARK
31	507.6	17.5	6828	12 ADQ19733	ADq19733 Human sof
32	507.6	17.5	6828	12 ADP43253	ADp43253 Human pro
33	507.6	17.5	6828	13 ADR25680	ADr25680 Breast ca
34	507.6	17.5	6854	12 ADQ23883	ADq23883 Human sof
35	506.4	17.5	2884	4 AAD14328	AAZ14328 Human pro
36	360.2	12.4	587	4 AAK93296	AAK93296 Human CDN
37	360.2	12.4	587	4 AAK91887	AAK91887 Human CDN
38	360.2	12.4	587	12 ADL29723	ADl29723 5' end of
39	360.2	12.4	587	12 ADL28314	ADl28314 5' end of
40	356.8	12.3	3594	5 AAD03994	AAZ03994 Human pro
41	296.8	10.2	1454	4 ABL21269	ABl21269 Drosophila
42	268.2	9.2	1723	4 AAH13802	AAh13802 Human CDN
43	236.6	8.2	1594	4 AAF44655	AAf44655 Novel pro
44	236.6	8.2	1594	8 AAF60326	AAf60326 Human 207
45	236.6	8.2	1594	12 ADI29353	ADI29353 Mouse MAR

ALIGNMENTS

RESULT 1
AAC90433
ID AAC90433 standard; cDNA; 2902 BP.
XX
AC AAC90433;
XX
DT 19-MAR-2001 (first entry)
XX
DE Murine lymph node Stromal cell kinase 1 coding sequence.
XX
KW Murine; lymph node Stromal cell kinase; Msk-1; autoimmune disorder;
KW wound healing; periodontal disease; inflammatory disease; tumour;
KW infection; allergy; ss.
XX
XX OS Mus musculus.
XX
XX WO200073468-A1.
XX
PD 07-DEC-2000.
XX
PE 26-MAY-2000; 2000WO-US014696.
XX
PR 28-MAY-1999; 99US-0136781P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Bird TR, Virca GD, Martin U, Anderson DM;
XX
XX WPI; 2001-061546/07.
XX
XX P-PSDB; AAB50056.
XX
PT Novel murine and human kinase nucleic acids useful for treating
PT inflammations, infections, tumors, allergies, autoimmune diseases, and
PT for stimulating or suppressing immune responsee.
XX
XX Claim 1; Page 86-87; 106pp; English.
XX
XX The present sequence is the coding sequence for Murine lymph node Stromal
XX cell kinase 1 (MSK-1). The protein encoded by the present sequence is
XX useful for treating a variety of disorders listed in the disclosure of
XX the specification, including autoimmune disorders, allergic reactions,
XX myeloid or lymphoid cell deficiencies, wound healing and tissue repair
XX and replacement, burns, incisions and ulcers, periodontal disease,

CC Inflammatory diseases, tumours and bacterial, viral or fungal infection
XX Sequence 2902 BP; 654 A; 794 C; 796 G; 658 T; 0 U; 0 Other;
Query Match 100.0%; Score 2902; DB 4; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACTAGTGTATCCAAAGATTCCGACAGAGGCTGTGCGGTGGCTGTGACTCTTGAG 60
DB 1 CACTAGTGTATCCAAAGATTCCGACAGAGGCTGTGCGGTGGCTGTGACTCTTGAG 60
QY 61 CCCGCGGCTCAGCGCGGCTGTGCTGTGCTGCGGACCACTCCCACTCCCTGCGGCTCCCGCA 120
DB 61 CCCGCGGCTCAGCGCGGCTGTGCTGTGCTGCGGACCACTCCCACTCCCTGCGGCTCCCGCA 120
QY 121 CCAATGAGTGTGGTGGCTTACTCAGGCGCCGAGCGAGGCTCCCTCGGCTCCGCTG 180
DB 121 CCAATGAGTGTGGTGGCTTACTCAGGCGCCGAGCGAGGCTCCCTCGGCTCCGCTG 180
QY 181 CTTGCGAGAGCGCCCGGCTGCGGCGGACGCGGCTCATCAAGTCCGCTTAACCTTGATGA 240
DB 181 CTTGCGAGAGCGCCCGGCTGCGGCGGACGCGGCTCATCAAGTCCGCTTAACCTTGATGA 240
QY 241 AGAAGCAGGCGGTGAAGGCGACCATGACAAACAACTGCGGCGACCGCTAGAGATTCC 300
DB 241 AGAAGCAGGCGGTGAAGGCGACCATGACAAACAACTGCGGCGACCGCTAGAGATTCC 300
QY 301 TGGAGAGCTGTGGCGAAGGCGACCTTACGGAAGGTGAAGAGGCGACGAGAGCTCGGCGC 360
DB 301 TGGAGAGCTGTGGCGAAGGCGACCTTACGGAAGGTGAAGAGGCGACGAGAGCTCGGCGC 360
QY 361 GTCTGTGGCTCATCAAGTCCATCAGAAAGCAAAATTCAAAGATGAGCAGGATGTGCTGC 420
DB 361 GTCTGTGGCTCATCAAGTCCATCAGAAAGCAAAATTCAAAGATGAGCAGGATGTGCTGC 420
QY 421 ACATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCACTATTCCTATCC 480
DB 421 ACATACGAGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCACTATTCCTATCC 480
QY 481 ATGAAGTGTGTGAAGATGAGCAAGATTGATGTTCATGAGATATGCCAGCGGAGCG 540
DB 481 ATGAAGTGTGTGAAGATGAGCAAGATTGATGTTCATGAGATATGCCAGCGGAGCG 540
QY 541 ATCTGATGATTTACATCAAGTGAAGCGGCGACGCTGAGTGAAGCGGAGCGGCAATTTCT 600
DB 541 ATCTGATGATTTACATCAAGTGAAGCGGCGACGCTGAGTGAAGCGGAGCGGCAATTTCT 600
QY 601 TCCGACGAGATCGTGTGCGGCTGCACTACGCGCACGGAACGGGATGTTTCAACGGAATC 660
DB 601 TCCGACGAGATCGTGTGCGGCTGCACTACGCGCACGGAACGGGATGTTTCAACGGAATC 660
QY 661 TCAAGCTGGAAGAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCC 720
DB 661 TCAAGCTGGAAGAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCC 720
QY 721 TCTTCAAACCTGTACCAAAAGGCAAGTTCTCCAGAGTTCGTGTGGAGGCCCTCTACG 780
DB 721 TCTTCAAACCTGTACCAAAAGGCAAGTTCTCCAGAGTTCGTGTGGAGGCCCTCTACG 780
QY 781 CTTGCGAGAGTGTGAAGGGAAGCCCTTATGTGGGCGGAGAGTGAAGAGTGTGCTC 840
DB 781 CTTGCGAGAGTGTGAAGGGAAGCCCTTATGTGGGCGGAGAGTGAAGAGTGTGCTC 840
QY 841 TGGGCGTTCCTCTGTATCATCTGTGTGATGACCATATGCCCTTTGACGGGCGAGATATA 900
DB 841 TGGGCGTTCCTCTGTATCATCTGTGTGATGACCATATGCCCTTTGACGGGCGAGATATA 900
QY 901 AAAACATGTGTGAAGCAAAATCAAGTAACGGGGCTTACCGTGAAGCCGCCAAGCGCTCCGATG 960
DB 901 AAAACATGTGTGAAGCAAAATCAAGTAACGGGGCTTACCGTGAAGCCGCCAAGCGCTCCGATG 960
QY 961 CTTGTGCGCTGATCCGGGTGGCTTTAATGTGTGAACCCCAACCGTGTGGGCCACATGTGAGG 1020

DB 961 CTTGTGCGCTGATCCGGGTGGCTTTAATGTGTGAACCCCAACCGTGTGGGCCACATGTGAGG 1020
QY 1021 ATGAGCCAGTCAATGTGTGGTCAACTGGGTTTACACACCGAGATGCGGGAACAGGAAG 1080
DB 1021 ATGAGCCAGTCAATGTGTGGTCAACTGGGTTTACACACCGAGATGCGGGAACAGGAAG 1080
QY 1081 CCTGTGAGAGGTGGACCTCATGTGTGACTTTGGCGGGCTCCATGCGGAGTGTGT 1140
DB 1081 CCTGTGAGAGGTGGACCTCATGTGTGACTTTGGCGGGCTCCATGCGGAGTGTGT 1140
QY 1141 TAGGTGCTCTGCGCGCCCTCCTGAGAGATGAGCAAGTGTGAGCTTCTTAAGC 1200
DB 1141 TAGGTGCTCTGCGCGCCCTCCTGAGAGATGAGCAAGTGTGAGCTTCTTAAGC 1200
QY 1201 AGCAGTGCAGGAGGTGAAGACCTGTACTGCGCTGAGAGCGGCAACTTCTTTAAGA 1260
DB 1201 AGCAGTGCAGGAGGTGAAGACCTGTACTGCGCTGAGAGCGGCAACTTCTTTAAGA 1260
QY 1261 AGTCCGGAAGAGATGACATGTCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCT 1320
DB 1261 AGTCCGGAAGAGATGACATGTCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCT 1320
QY 1321 CTTCTGCGCCCTGGCAAGAGAGCCTTAAGCTTCGAAAGGCAATTCTCAAGAAAAGTCT 1380
DB 1321 CTTCTGCGCCCTGGCAAGAGAGCCTTAAGCTTCGAAAGGCAATTCTCAAGAAAAGTCT 1380
QY 1381 CTACCTGTGACAGGAGAGTACAGAGAGACCTCAGAACTCAAGCCGAGCTGTATCTC 1440
DB 1381 CTACCTGTGACAGGAGAGTACAGAGAGACCTCAGAACTCAAGCCGAGCTGTATCTC 1440
QY 1441 CAGGCGAGCTGTCCCTGTGTATCCCTGCTCCCAAGAAAGGATCTTAAGAGTCTC 1500
DB 1441 CAGGCGAGCTGTCCCTGTGTATCCCTGCTCCCAAGAAAGGATCTTAAGAGTCTC 1500
QY 1501 GACAGGATGATGTGGTTACTACTCTCTCAAGAGCCAGAGCTGTGGGAACTCTAG 1560
DB 1501 GACAGGATGATGTGGTTACTACTCTCTCAAGAGCCAGAGCTGTGGGAACTCTAG 1560
QY 1561 ACAGCAGTGAATGTGTGTGAGTGGGAGCCCGTGAAGCAAGTCTCCACAGGCTTCAAG 1620
DB 1561 ACAGCAGTGAATGTGTGTGAGTGGGAGCCCGTGAAGCAAGTCTCCACAGGCTTCAAG 1620
QY 1621 GGCCTCTCTCCACCGCAGAGGCAATCTCAATGCGAAGTTCTCCGCAAGGCT 1680
DB 1621 GGCCTCTCTCCACCGCAGAGGCAATCTCAATGCGAAGTTCTCCGCAAGGCT 1680
QY 1681 TAGAAGCATTACCCCTAGACCTTGGCTCCCTGAGCAAACTGAGGCTCTCCCATCTGT 1740
DB 1681 TAGAAGCATTACCCCTAGACCTTGGCTCCCTGAGCAAACTGAGGCTCTCCCATCTGT 1740
QY 1741 CAGCCCGGCGCAGCCGCTCAGAGGCGTGTGAGTGAAGACAGATCTCTCCGAGT 1800
DB 1741 CAGCCCGGCGCAGCCGCTCAGAGGCGTGTGAGTGAAGACAGATCTCTCCGAGT 1800
QY 1801 CTTTGAACAATTGAACCTTGTGAGAGTCTTCCGAAACCCCACTGAGGGGCTGTGTGT 1860
DB 1801 CTTTGAACAATTGAACCTTGTGAGAGTCTTCCGAAACCCCACTGAGGGGCTGTGTGT 1860
QY 1861 CTGTGACAACTGAGGGGGCTTGAAGAGCTCTCCCTGAGAAAGTCTGAAACGATGTGTGC 1920
DB 1861 CTGTGACAACTGAGGGGGCTTGAAGAGCTCTCCCTGAGAAAGTCTGAAACGATGTGTGC 1920
QY 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTGTGACAGCTGCCAAGAGGTGACTGACGCT 1980
DB 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTGTGACAGCTGCCAAGAGGTGACTGACGCT 1980
QY 1981 ACAAGCAAGCCTTAGAATGTCTCAAGTCAAGTGAAGAGAGAGATGTGCTTAGT 2040
DB 1981 ACAAGCAAGCCTTAGAATGTCTCAAGTCAAGTGAAGAGAGAGATGTGCTTAGT 2040
QY 2041 ATGGGGTAGGCTGTGAGAGGTTTTCAGAGAAACCTGTGGGTGGATTCCTCAAGTGAATA 2100

Db 2041 ATGGGTTAGGCTCTGAGAGGGTTTGACAGAGAACCTTGGGATTCCTCAGTAATA 2100
 Qy 2101 GAGTACATCAAGGGCTCTACGCTGACGCTGACCTGAACCTGAAGATGAGAAATGCG 2160
 Db 2101 GAGTACATCAAGGGCTCTACGCTGACGCTGACCTGAACCTGAAGATGAGAAATGCG 2160
 Qy 2161 ATTGATGTGAAAGAAATGAGAAACCTTGTCTGCCGAGTGTATATAGTGGGCTGCGA 2220
 Db 2161 ATTGATGTGAAAGAAATGAGAAACCTTGTCTGCCGAGTGTATATAGTGGGCTGCGA 2220
 Qy 2221 GGTGCTCACTCTCTTGTGCAATGAGTGAACCCATGACATTTCCACCCCTGTTCTGCG 2280
 Db 2221 GGTGCTCACTCTCTTGTGCAATGAGTGAACCCATGACATTTCCACCCCTGTTCTGCG 2280
 Qy 2281 CTGCACTTCAATGATTTCTGTTTCAATCAACCAAGGATGAGAACTGACTTCT 2340
 Db 2281 CTGCACTTCAATGATTTCTGTTTCAATCAACCAAGGATGAGAACTGACTTCT 2340
 Qy 2341 GGGAGGTATGTGTAGTACTGCTCCATTTATAGAGAGAAACAGCTCTGGTTTCAATCT 2400
 Db 2341 GGGAGGTATGTGTAGTACTGCTCCATTTATAGAGAGAAACAGCTCTGGTTTCAATCT 2400
 Qy 2401 CTGCTGTGTGATCTTCAAAAGACTGGGAAAGACTGGACCGCTGTGACTTCAATCA 2460
 Db 2401 CTGCTGTGTGATCTTCAAAAGACTGGGAAAGACTGGACCGCTGTGACTTCAATCA 2460
 Qy 2461 GGGGACCAAGATGCCCTGAGACCCCATCTTATGATCTGAGAGCTTGAAGCTGTT 2520
 Db 2461 GGGGACCAAGATGCCCTGAGACCCCATCTTATGATCTGAGAGCTTGAAGCTGTT 2520
 Qy 2521 CCTAGTACCAAGATGTGATGATGATCTGTTTCTAGGCGCAAGGACCTAGATGTGC 2580
 Db 2521 CCTAGTACCAAGATGTGATGATGATCTGTTTCTAGGCGCAAGGACCTAGATGTGC 2580
 Qy 2581 TGACTATTTATTTTGTGATCTGACTCTGTTTGTGTTTGTGTTTGTGTTG 2640
 Db 2581 TGACTATTTATTTTGTGATCTGACTCTGTTTGTGTTTGTGTTTGTGTTG 2640
 Qy 2641 TTTTGTGTTTAAATGTAATTTGCTGCTTCAATATATGTAATGTAATGTAATGTAATG 2700
 Db 2641 TTTTGTGTTTAAATGTAATTTGCTGCTTCAATATATGTAATGTAATGTAATGTAATG 2700
 Qy 2701 TCCACTGTGCTACATGATTTATGATACAGAAATATTTGGCAATGATGTCCTCTATTC 2760
 Db 2701 TCCACTGTGCTACATGATTTATGATACAGAAATATTTGGCAATGATGTCCTCTATTC 2760
 Qy 2761 AAGGGGGGGGGGGGGCTTTTCAATATGATGTTGAGCACTGTGTGATGATGATGATG 2820
 Db 2761 AAGGGGGGGGGGGGGCTTTTCAATATGATGTTGAGCACTGTGTGATGATGATGATG 2820
 Qy 2821 GTGCTTCAACCAAGCTGCGCAACCTCTCTCAATCTTCAATCTGATGAGTGTGCA 2880
 Db 2821 GTGCTTCAACCAAGCTGCGCAACCTCTCTCAATCTTCAATCTGATGAGTGTGCA 2880
 Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
 Db 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
 RESULT 2
 ID ABD33083 standard; cDNA; 3073 BP.
 AC ABD33083;
 XX 18-NOV-2004 (first entry)
 DE Murine cancer-associated (CA) cDNA MR07-004.
 XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW B6; cancer; cytostatic.
 OS Mus musculus.

XX XH
 PN W02004058146-A2.
 XX 15-JUL-2004.
 PD 15-DEC-2003; 2003MO-US040081.
 XX 15-DEC-2003; 2003MO-US040081.
 PF 17-DEC-2002; 2002JUS-00322281.
 PR (SAGR-) SAGRES DISCOVERY INC.
 PA Morrie DW, Malandro MS;
 PI WPI; 2004-499109/47.
 DR P-PSDB; AB084756.
 XX
 PT Novel human cancer associated protein encoded within open reading frame
 of cancer associated gene, useful as targets for diagnosing cancer.
 PS Disclosure; SEQ ID NO 22; 182bp; English.
 CC The invention relates to cancer-associated proteins (CAP) and the cancer-
 associated (CA) nucleic acids encoding them. The invention also relates
 to a method for treating cancers involving administering to a patient an
 inhibitor of CAP, and a method of screening for anticancer activity in a
 potential drug involving providing a cell that expresses a CA gene,
 contacting a tissue sample derived from a cancer cell with an anticancer
 drug candidate and monitoring the effect of the anticancer drug candidate
 on expression of the CA gene. The CAP proteins are useful for detecting
 cancer associated with expression of a CAP protein in a test cell sample
 and for screening for a bioactive agent capable of modulating the
 activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 cancer, involving determining the expression of a CA nucleic acid in a
 tissue. This sequence represents murine CA cDNA of the invention. Note:
 The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at http://wipo.int/pub/published_pcl_sequences
 XX
 SQ Sequence 3073 BP; 690 A; 827 C; 844 G; 712 T; 0 U; 0 Other;
 Query Match 98.7%; Score 2863.4; DB 13; Length 3073;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2864; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 32 CGTGTGGGGTGGGGCTGTGACCTGTGAGCCCGGCTGAGCGGCTGACTGCTGCG 91
 Db 1 CGTGTGGGGTGGGGCTGTGACCTGTGAGCCCGGCTGAGCGGCTGACTGCTGCG 60
 Qy 92 CCGAACCACCTCACTTCGCGGTCCCGCACATGATGATGATGATGATGATGATGATG 151
 Db 61 CCGAACCACCTCACTTCGCGGTCCCGCACATGATGATGATGATGATGATGATGATG 120
 Qy 152 GAGCCAGGCTTCCTCGGCTTCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 211
 Db 121 GAGCCAGGCTTCCTCGGCTTCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 180
 Qy 212 GCTCATTAATGCTCTTAACTCTGATGAAAGACGCGGCTGAAAGGCGACCATCA 271
 Db 181 GCTCATTAATGCTCTTAACTCTGATGAAAGACGCGGCTGAAAGGCGACCATCA 240
 Qy 272 ACACAACTCGGCAACCGTACGATCTCTGAGACGCTGAGCAAGGCGCACTTACGCGGAA 331
 Db 241 ACACAACTCGGCAACCGTACGATCTCTGAGACGCTGAGCAAGGCGCACTTACGCGGAA 300
 Qy 332 GGTGAAGAGGCAAGAGAGCTTCGGGGCTGTGGTGGCCATCAATGCTCAAGGAAAGA 391
 Db 301 GGTGAAGAGGCAAGAGAGCTTCGGGGCTGTGGTGGCCATCAATGCTCAAGGAAAGA 360
 Qy 392 CAATAATCAAGATGAGAGATCTGTGCAATACGAGAGGAGGATTAAGATCAATGCTTCC 451
 Db 361 CAATAATCAAGATGAGAGATCTGTGCAATACGAGAGGAGGATTAAGATCAATGCTTCC 420
 Qy 452 ACTCAACCAACCCCACTATTCATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 511

```
Db 421 ACTCAACCAACCCCACTCATCTTCCATCATGATGTTTGAGAAATGACGAAGATTGT 480
Qy 512 GATTGTGATGAGATATCCAGCCGAGCGATCTGTATGATTAATCATCATGAGCGGCAAG 571
Db 481 GATTGTGATGAGATATCCAGCCGAGCGATCTGTATGATTAATCATCATGAGCGGCAAG 540
Qy 572 GCTGAGTGAACGGGAGCCGAGGCACTTTCTTCCGACAAATGATGTGTCTGCTGCACTG 631
Db 541 GCTGAGTGAACGGGAGCCGAGGCACTTTCTTCCGACAAATGATGTGTCTGCTGCACTG 600
Qy 632 CCAACGAGAAGGAGCTTCAACCGAGATCTCAAGCTGGAATAATCTCTTGAATGCA 691
Db 601 CCAACGAGAAGGAGCTTCAACCGAGATCTCAAGCTGGAATAATCTCTTGAATGCA 660
Qy 692 TGGAAAATCAAGATTGCTGAATTTGGCTCTTCCAACTGTACCAAGGCAAGTTCT 751
Db 661 TGGAAAATCAAGATTGCTGAATTTGGCTCTTCCAACTGTACCAAGGCAAGTTCT 720
Qy 752 CCAAGAGCTTCTGTGGAGGCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCTA 811
Db 721 CCAAGAGCTTCTGTGGAGGCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCTA 780
Qy 812 TGTGGAGCCCAAGAGTGAAGAGCTGTCTGTGGGCTTCTCTGTATCAATCTGTGTGATG 871
Db 781 TGTGGAGCCCAAGAGTGAAGAGCTGTCTGTGGGCTTCTCTGTATCAATCTGTGTGATG 840
Qy 872 CACCATGCTCTTTGACGGGAGAGATCAATAAAACATGATGAAGCAATCAAGTACGGG 931
Db 841 CACCATGCTCTTTGACGGGAGAGATCAATAAAACATGATGAAGCAATCAAGTACGGG 900
Qy 932 TTACCGTGAACCCGCTCCAGCCGCTCCGATGCTGTGGCTGTATCCGCTGTGCTTAAATG 991
Db 901 TTACCGTGAACCCGCTCCAGCCGCTCCGATGCTGTGGCTGTATCCGCTGTGCTTAAATG 960
Qy 992 GAACCCCAACCCGCTCCGAGCCCACTGAGAGATGTAGCCAGTCAATGTGTGGCTCAACTG 1051
Db 961 GAACCCCAACCCGCTCCGAGCCCACTGAGAGATGTAGCCAGTCAATGTGTGGCTCAACTG 1020
Qy 1052 TTACACCAACCGAGTGGGGAACAGAAAGCTTGCATGAGGTGGGACCCCTACTGTATG 1111
Db 1021 TTACACCAACCGAGTGGGGAACAGAAAGCTTGCATGAGGTGGGACCCCTACTGTATG 1080
Qy 1112 CTTTGGCCGGGCTCTCAATGCGAGATGTATACGTGCTCTCTCGGCCCTCTCTGAGAA 1171
Db 1081 CTTTGGCCGGGCTCTCAATGCGAGATGTATACGTGCTCTCTCGGCCCTCTCTGAGAA 1140
Qy 1172 TGGAGCCAGGTGTGACGCTTCTTCAAGCAGACATGTCCGGGAGGTGGAAGCACTATCC 1231
Db 1141 TGGAGCCAGGTGTGACGCTTCTTCAAGCAGACATGTCCGGGAGGTGGAAGCACTATCC 1200
Qy 1232 TGGGCTGAGACGGGCAATTTCTTTAAAGATCCCGAAAGAGATGACATGTGCTCAAA 1291
Db 1201 TGGGCTGAGACGGGCAATTTCTTTAAAGATCCCGAAAGAGATGACATGTGCTCAAA 1260
Qy 1292 TCTGCAAGGTGATCCCGGCTGAGATACCTTCTTCCGCTGCGCAAGGACCTTTAAGCT 1351
Db 1261 TCTGCAAGGTGATCCCGGCTGAGATACCTTCTTCCGCTGCGCAAGGACCTTTAAGCT 1320
Qy 1352 TCCGAAGAGGATTTCTCAAGAAAGTCTACTCTGTACGGGAGGATCAAGAGGAGCC 1411
Db 1321 TCCGAAGAGGATTTCTCAAGAAAGTCTACTCTGTACGGGAGGATCAAGAGGAGCC 1380
Qy 1412 TCAAGAACTGAGACCGGTGCTGATACTTCCAGGAGACCTGTCCCTGTATCCCTGT 1471
Db 1381 TCAAGAACTGAGACCGGTGCTGATACTTCCAGGAGACCTGTCCCTGTATCCCTGT 1440
Qy 1472 CCCAAGGAAGGATCTTTAAAGATCTCGACAGCTGATGTGTTATCTCTCTCC 1531
Db 1441 CCCAAGGAAGGATCTTTAAAGATCTCGACAGCTGATGTGTTATCTCTCTCC 1500
Qy 1532 AGAGCCAGGAGATCTGGGGAACCTTGAAGCAGAGATGTGTTTGTGATGGGAGCC 1591
Db 1501 AGAGCCAGGAGATCTGGGGAACCTTGAAGCAGAGATGTGTTTGTGATGGGAGCC 1560
Qy 1592 CGTGAAGCAGAATCTTCCACAGGCTTCAAGGCTCTCTTCCACCGCAAGGCAATTTCAA 1651
Db 1561 CGTGAAGCAGAATCTTCCACAGGCTTCAAGGCTCTCTTCCACCGCAAGGCAATTTCAA 1620
Qy 1652 ACTCAATGGAATCTTCCCGGACAGCTTAAAGAGCACTAACCTTAGACCTTTGGCTC 1711
Db 1621 ACTCAATGGAATCTTCCCGGACAGCTTAAAGAGCACTAACCTTAGACCTTTGGCTC 1680
Qy 1712 CTTGAGCAACATGAGCTCTCCATCTCCATCTGTGAGCCGAGCCGCTCAGGGGCTGT 1771
Db 1681 CTTGAGCAACATGAGCTCTCCATCTCCATCTGTGAGCCGAGCCGCTCAGGGGCTGT 1740
Qy 1772 GAGTGAAGACAGATCTGTCTCTCGAGTCTTTGAACAAATGGAATCTGTCTGAAGCT 1831
Db 1741 GAGTGAAGACAGATCTGTCTCTCGAGTCTTTGAACAAATGGAATCTGTCTGAAGCT 1800
Qy 1832 TCCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACTGAGGGGCTTGAAGCACC 1891
Db 1801 TCCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACTGAGGGGCTTGAAGCACC 1860
Qy 1892 TCCCTGAGAAAGTCTGAAGGATGTGTGACAGAAATCCTTGGGGATATGCTGCTTTCT 1951
Db 1861 TCCCTGAGAAAGTCTGAAGGATGTGTGACAGAAATCCTTGGGGATATGCTGCTTTCT 1920
Qy 1952 GACAGACTGCCAAGAGTGAATGCAAGCTTACAGCAAGCCCTTGAATCTGTCAAGCT 2011
Db 1921 GACAGACTGCCAAGAGTGAATGCAAGCTTACAGCAAGCCCTTGAATCTGTCAAGCT 1980
Qy 2012 CAGCTGAGGAAGGAGATGTGTGCTGTATGTGGGATAGGCTGAGAGGGTTTGCAGAG 2071
Db 1981 CAGCTGAGGAAGGAGATGTGTGCTGTATGTGGGATAGGCTGAGAGGGTTTGCAGAG 2040
Qy 2072 AACCTGGGATCGGATTTCTCTCAAGTAAATAGATACATCAAGGCTCTACCTGACGCT 2131
Db 2041 AACCTGGGATCGGATTTCTCTCAAGTAAATAGATACATCAAGGCTCTACCTGACGCT 2100
Qy 2132 GACTGAACTTGAAAGATGAGAAATGCAATGTATGTGAAAGAAATGGAAACCTTTGCT 2191
Db 2101 GACTGAACTTGAAAGATGAGAAATGCAATGTATGTGAAAGAAATGGAAACCTTTGCT 2160
Qy 2192 GCCCGAGTGTATAGTGGGTGGCTGGAAGTGTCTACCTCTTTGTGCAATGAGTCA 2251
Db 2161 GCCCGAGTGTATAGTGGGTGGCTGGAAGTGTCTACCTCTTTGTGCAATGAGTCA 2220
Qy 2252 CCCATGACATTTCCCAACCCCTGTTCTCTGAGCTGACCTTCAATAGTTCTGTTCCATC 2311
Db 2221 CCCATGACATTTCCCAACCCCTGTTCTCTGAGCTGACCTTCAATAGTTCTGTTCCATC 2280
Qy 2312 AACCAACGAGGTTAGAACCTTGAATCTTCTGGAAGGTAAATGTATGATGATCGCAATAT 2371
Db 2281 AACCAACGAGGTTAGAACCTTGAATCTTCTGGAAGGTAAATGTATGATGATCGCAATAT 2340
Qy 2372 AGAAGGAACACGCTCTGATTTCCATCTGTGCTGTGTGATCTCAAAACCTGGGAAG 2431
Db 2341 AGAAGGAACACGCTCTGATTTCCATCTGTGCTGTGTGATCTCAAAACCTGGGAAG 2400
Qy 2432 ACTCGAACGCTGTTGATCTTCAATCAAGGGAACAGATGACCCCTGAGACCCCATCTTAA 2491
Db 2401 ACTCGAACGCTGTTGATCTTCAATCAAGGGAACAGATGACCCCTGAGACCCCATCTTAA 2460
Qy 2492 ATCTCAGAGAATTTGAACCTTGAAGCTGTTCTTATGATCCAGATGTGATGATGCTGT 2551
Db 2461 ATCTCAGAGAATTTGAACCTTGAAGCTGTTCTTATGATCCAGATGTGATGATGCTGT 2520
Qy 2552 TTCTCAGGCCAAGGAGCTTGAATGTGCTGAATTTATTTTGTGATTTCTCATTT 2611
Db 2521 TTCTCAGGCCAAGGAGCTTGAATGTGCTGAATTTATTTTGTGATTTCTCATTT 2580
Qy 2612 CTGTTTTTGTGTTTTTGTGTTTTGTTTTGTTTTTAAAGATTTTGTGCTTTC 2671
Db 2581 CTGTTTTTGTGTTTTTGTGTTTTGTTTTGTTTTTAAAGATTTTGTGCTTTC 2640
```

QY 2672 AATAATGTAATGCTGTCTTCTGAGAACTCCACTGTCCTGAAAGTTATGACAGAG 2731
 CC 2641 AATAATGTAATGCTGTCTTCTGAGAACTCCACTGTCCTGAAAGTTATGACAGAG 2700
 Db 2732 AAGATATTTGGGAATGATCCCTCTATTCAGAGGGGGGGGGGGGGCTTTTCAATATGATG 2791
 QY 2701 AAGATATTTGGGAATGATCCCTCTATTCAGAGGGGGGGGGGGGGCTTTTCAATATGATG 2760
 Db 2792 TCTTGAGCACTGTCTGATGATGCTCCAGTCCCTTCACACCCAGGCTGGCCACCTCC 2851
 QY 2761 TCTTGAGCACTGTCTGATGATGCTCCAGTCCCTTCACACCCAGGCTGGCCACCTCC 2820
 Db 2852 CTCATCTTTCATCTGTGGCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 2896
 QY 2821 CTCATCTTTCATCTGTGGCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 2865
 Db
 RESULT 3
 AAD31710
 ID AAD31710 standard; cDNA; 2929 BP.
 XX AAD31710;
 AC
 XX 18-JUN-2002 (first entry)
 DT
 XX
 DE Rat SNF1/AMPK-Related Kinase (SNARK) cDNA.
 XX
 KM Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;
 KM glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;
 KM hyperglycaemic; drug screening; hypoglycaemia; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS /*tag= a
 FT 83..1975
 FT /product= "Rat SNARK protein"
 XX
 PN W0200212456-A2.
 PD 14-FEB-2002.
 XX
 PF 02-AUG-2001; 2001WO-CA001109.
 XX
 PR 03-AUG-2000; 2000US-0222650P.
 PR 12-MAR-2001; 2001US-0274613P.
 PR 28-MAR-2001; 2001CA-02340783.
 XX
 PA (ONBO-) 1149336 ONTARIO INC.
 XX
 PI Drucker DJ, Rosen CF, Lefebvre DL;
 XX
 DR WPI; 2002-241747/29.
 DR P-PSDB; AAE19885.
 XX
 PT AMPK (AMP-activated protein kinase)-related kinase, designated SNARK
 PT polypeptides and polynucleotides, useful for treating or preventing
 PT diabetes, or other disorders of lipoprotein production leading to
 PT increased levels of cholesterol.
 XX
 PS Example; Fig 2; 94pp; English.
 XX
 CC The invention relates to an AMPK (AMP-activated protein kinase)-related
 CC kinase, designated SNARK polypeptides and polynucleotides. SNARK
 CC (SNF1/AMP-activated protein kinase) is involved in stress response to
 CC glucose deprivation. The polynucleotides are useful for expressing SNARK
 CC protein in isolated form or as a protein conjugate. Activation of SNARK
 CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
 CC in other cell types such as heart and skeletal muscles, as well as
 CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
 CC to have insulin-like effects that would enhance the disposal of glucose
 CC into muscle and reduce plasma glucose for the treatment of diabetes and

CC some type of disorders of lipoprotein production leading to increased
 CC levels of cholesterol or triglycerides. SNARK or its variants may be
 CC administered to a subject to treat or prevent a disease associated with
 CC decreased expression of SNARK, such as diabetes. SNARK antibodies are
 CC used to modulate SNARK activity either in vivo for therapeutic purposes,
 CC or in vitro, for drug screening and related investigational purposes.
 CC SNARK antagonists may be administered to increase fuel production,
 CC decrease glucose uptake and increase levels of blood glucose in a patient
 CC suffering from hypoglycaemia. The present sequence is rat SNARK cDNA
 CC
 SQ Sequence 2929 BP; 650 A; 788 C; 791 G; 700 T; 0 U; 0 Other;
 Query Match 71.2%; Score 2066.8; DB 6; Length 2929;
 Best Local Similarity 86.4%; Pred. No. 0;
 Matches 2488; Conservative 0; Mismatches 277; Indels 115; Gaps 14;
 QY 49 GTGACCTTGTAGCCCGGCTCAACGCGCGCTGTCTACTGCTGCGGACCCCTCCACTC 108
 Db 9 GTGACCTTGTAGCCCGGCTCTCCGCGCTGTCTGCTGCTGCGGACCCCTCCGCTC 68
 QY 109 GCGGTCCCGGACCATGAGTCGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCG 168
 Db 69 GCGGTCCCGGACCATGAGTCGGTGGCTTACTCCAGCGCGGGAACCTGCTCCCTCG 128
 QY 169 CCTCCGCGCTGGCTCTGAGAGCGCCCGCGCTGGCGGACGGCTCATCAAGTCGCTA 228
 Db 129 CCTCCGCGCTGGCGGACCATGAGTCGGTGGCTTACTCCAGCGCGGGAACCTGCTCC 188
 QY 229 AACCTGTATGTAAGAAGCGCGGTGAAAGCGGACCATCAACAAACAACTGCGGCA 288
 Db 189 AACCTGTATGTAAGAAGCGCGGTGAAAGCGGACCATCAACAAACAACTGAGGCA 248
 QY 289 GCTACGAGTTCTGTGAGAGCGCTGGGCAAGGSCACTACGGGAAGGTGAAGAGCAG 348
 Db 249 GCTACGAGTTCTGTGAGAGCGCTGGGCAAGGSCACTACGGGAAGGTGAAGAGCAG 308
 QY 349 AGACCTCGGCGCTGTGTGCGCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGC 408
 Db 309 AGACCTCGGCGCGCTGTGTGCGCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGC 368
 QY 409 AGATCTGTGCAATACGAGGAGATTGATGATGATCTTCACTCAACCACTCCCA 468
 Db 369 AGATCTGTGCAATACGAGGAGATTGATGATGATCTTCACTCAACCACTCCCA 428
 QY 469 TCATTGGCATCAGTAAGTGTGTAAGATGACCAAGATTGTGATGATGATGATG 528
 Db 429 TCATTGGCATCAGTAAGTGTGTAAGATGACCAAGATTGTGATGATGATGATG 488
 QY 529 CCAGCCGAGGCGATCTGTATGATTCATCAGTGAAGCGGCGCAAGCTGAGTGAAGCG 588
 Db 489 CCAGCCGAGGCGATCTGTATGATTCATCAGTGAAGCGGCGCAAGCTGAGTGAAGCG 548
 QY 589 CCAGGCAATTTCTTCCGACAGATGTGTCTGCGCTGCACTACTGCGACCAAGAGGATG 648
 Db 549 CCAGGCAATTTCTTCCGACAGATGTGTCTGCGCTGCACTACTGCGACCAAGAGGATG 608
 QY 649 TTCAACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCAATGGAAGAAATCAAGATTG 708
 Db 609 TTCAACCGAGATCTCAAGCTGGAAGAAATCTCTTCAATGCAATGGAAGAAATCAAGATTG 668
 QY 709 CTGACTTTGGGCTCTCCAACTGTACCAAGGCAAGGCTCTCCAGAGCTTCTGTGGA 768
 Db 669 CTGACTTTGGGCTCTCCAACTGTATCAAGGCAAGGCTCTCCAGAGCTTCTGTGGA 728
 QY 769 GCGCTTCTTACGCTGCTGTAGATGATCAACGGAAGCCCTATGTGGGCCCAAGGTG 828
 Db 729 GCGCTTCTTACGCTGCTGTAGATGATCAACGGAAGCCCTATGTGGGCCCAAGGTG 788
 QY 829 ACAGCTGTCTGTGGGGGTTCTCTGTATCAATCTGTGTGATGAGCAATGCTTTGAG 888
 Db 789 ACAGCTGTCTGTGGGGGTTCTCTGTATCAATCTGTGTGATGAGCAATGCTTTGAG 848
 QY 889 GGCAGGATCATTAACAACCTGTGAGCAAAATCAAGTAAACGGGGCTTACCGTGAAGCCGCA 948

KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW de; cancer; cytostatic.
XX Mus musculus.
XX
XX NO2004058146-A2.
PN
XX
XX 15-JUL-2004.
PD
XX
XX 15-DEC-2003; 2003WO-US040081.
PF
XX
XX 17-DEC-2002; 2002US-00322281.
PR
XX
XX (SAGR-) SAGRES DISCOVERY INC.
PA
XX
XX Morrie DW, Malandro MS;
PI
XX
XX WPI; 2004-499109/47.
DR
XX
XX
XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
XX
PS Disclosure; SEQ ID NO 21; 182pp; English.
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 37278 BP; 8909 A; 9465 C; 8785 G; 9687 T; 0 U; 432 Other;
Query Match 66.8%; Score 1938.8; DB 13; Length 37278;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 955 CCGATGCTGTGCTGCTGATCCGATGCTGTTAATGTGAACCCGCGTGGGCAAC 1014
DB 25129 CAGATGCTGTGCTGCTGATCCGATGCTGTTAATGTGAACCCGCGTGGGCAAC 25188
QY 1015 TGGAGATGTAGCCAGTCATTTGTGGTCAACTGGGGTTACACACCGGAGTCGGGAAC 1074
DB 25189 TGGAGATGTAGCCAGTCATTTGTGGTCAACTGGGGTTACACACCGGAGTCGGGAAC 25248
QY 1075 AGGAAGCCGTCGCGAGAGGTGGGCAACCTAGTGTGACTTTGGCGGCTCCACATGGGG 1134
DB 25249 AGGAAGCCGTCGCGAGAGGTGGGCAACCTAGTGTGACTTTGGCGGCTCCACATGGGG 25308
QY 1135 ACTGTTACGTGCTGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGTGACGTTCT 1194
DB 25309 ACTGTTACGTGCTGCTCTCGGCGCCCTCTCTGAGAAATGAGCAAGGTGTGACGTTCT 25368
QY 1195 TCAAGCAGCAGTGTGCGGAGGTGAAGCACTGTACTGGGTGAGCGGCAACATTTCTC 1254
DB 25369 TCAAGCAGCAGTGTGCGGAGGTGAAGCACTGTACTGGGTGAGCGGCAACATTTCTC 25428
QY 1255 TTAGAAGTCCCGAAGAGAGATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGG 1314
DB 25429 TTAGAAGTCCCGAAGAGAGATGACATGCTCAAAATCTGCAAGGTGACCGGCTGAGG 25488
QY 1315 ATACTCTTCTGCTGCTGCGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTCTCAAGAAA 1374

DB 25489 ATACTCTTCTGCTGCTGCGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTCTCAAGAAA 25548
QY 1375 AGTCTCTACCTCTGTCAGGAGAGGTATAGAGAGACCTCTCAGAACTCAGACCGGTGCTG 1434
DB 25549 AGTCTCTACCTCTGTCAGGAGAGGTATAGAGAGACCTCTCAGAACTCAGACCGGTGCTG 25608
QY 1435 ATACTCAGGAGAGCTGTGCTGCTGATTCCTGCTCCAGAGAAAGGCAATCTTAAAGA 1494
DB 25609 ATACTCAGGAGAGCTGTGCTGCTGATTCCTGCTCCAGAGAAAGGCAATCTTAAAGA 25668
QY 1495 AGTCTCAGAGCGGGAATCTGTGTTAATCTCTCTCCAGAGCCGAGAGTCTGGGAGAC 1554
DB 25669 AGTCTCAGAGCGGGAATCTGTGTTAATCTCTCTCCAGAGCCGAGAGTCTGGGAGAC 25728
QY 1555 TCTTAGAGCCAGATGATGTGTTGTGATGGGAGACCCGCTGAGACAGAACTCTCCACAG 1614
DB 25729 TCTTAGAGCCAGATGATGTGTTGTGATGGGAGACCCGCTGAGACAGAACTCTCCACAG 25788
QY 1615 CTTGAGGAGCTCTCTCTCCAGAGGAGGCAATTTCTCAAACTCATATGACAGTTCTCCGCA 1674
DB 25789 CTTGAGGAGCTCTCTCTCCAGAGGAGGCAATTTCTCAAACTCATATGACAGTTCTCCGCA 25848
QY 1675 CAGCTTAAAGAGCACTAACCTTAGCACTTTGGCTCCCTGAGACCACTGGCTCTCTCC 1734
DB 25849 CAGCTTAAAGAGCACTAACCTTAGCACTTTGGCTCCCTGAGACCACTGGCTCTCTCC 25908
QY 1735 ATCTGAGAGCCGAGCCGAGCCGCTCAGGAGGCTGTGAGTGAAGAGACAGATCTGCTCT 1794
DB 25909 ATCTGAGAGCCGAGCCGAGCCGCTCAGGAGGCTGTGAGTGAAGAGACAGATCTGCTCT 25968
QY 1795 CCGAGTCTTTGACCAATTTGACCTTGCTGACACCTCTCCGAAACCCCACTGAGGAGCT 1854
DB 25969 CCGAGTCTTTGACCAATTTGACCTTGCTGACACCTCTCCGAAACCCCACTGAGGAGCT 26028
QY 1855 GTGTGTCTGTGACCACTGAGGAGGCTTTGAGACGCTCTCTCAGAAAGTCTGAAGCAT 1914
DB 26029 GTGTGTCTGTGACCACTGAGGAGGCTTTGAGACGCTCTCTCAGAAAGTCTGAAGCAT 26088
QY 1915 GTGTGAGGAATCTGTGGGGATGAGCTTTTCTCTGACAGACTGCGCAAGAGGTGACTG 1974
DB 26089 GTGTGAGGAATCTGTGGGGATGAGCTTTTCTCTGACAGACTGCGCAAGAGGTGACTG 26148
QY 1975 CAGCTTACAGCAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGAGAGAGATGTGTC 2034
DB 26149 CAGCTTACAGCAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGAGAGAGATGTGTC 26208
QY 2035 CTTAGTATGGGGTATGCTCTGAGAGGGTTTGCAGAGAACTCTGGGTGCAATTCCTCAG 2094
DB 26209 CTTAGTATGGGGTATGCTCTGAGAGGGTTTGCAGAGAACTCTGGGTGCAATTCCTCAG 26268
QY 2095 TGAATAGATACATCAAGGGCTCTACGTCTGACCTGACCTGAACCTGAAAGATGAGAGA 2154
DB 26269 TGAATAGATACATCAAGGGCTCTACGTCTGACCTGACCTGAACCTGAAAGATGAGAGA 26328
QY 2155 AATGCAATGATGTGAAAGGAATGGGAACCTTGTGCTGCCGAGTGTATATGTGGGGTGG 2214
DB 26329 AATGCAATGATGTGAAAGGAATGGGAACCTTGTGCTGCCGAGTGTATATGTGGGGTGG 26388
QY 2215 CTTGAGAGGTGCTACCTCTTGTGTCATGATGTGTCACCAATGACATTTTCCACCTGTT 2274
DB 26389 CTTGAGAGGTGCTACCTCTTGTGTCATGATGTGTCACCAATGACATTTTCCACCTGTT 26448
QY 2275 CTGTGCTGACCTCTTACATTAAGTTTCTGTTTCAATCAACCAAGGTTTGAACCTCTGA 2334
DB 26449 CTGTGCTGACCTCTTACATTAAGTTTCTGTTTCAATCAACCAAGGTTTGAACCTCTGA 26508
QY 2335 CTTCTCGGAGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394
DB 26509 CTTCTCGGAGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26568
QY 2395 CCAATCTGCTGCTGTGATCTCAAAAGACTGTGGAGACTGGAACCGCTGTTTGAATTTCA 2454
DB 26569 CCAATCTGCTGCTGTGATCTCAAAAGACTGTGGAGACTGGAACCGCTGTTTGAATTTCA 26628

QY 2455 TCTCAAGGAGCAGATGCCCCGACCCCACTTTAGTCTCAGAGACTTGACCTTGA 2514
DB 26629 TCTCAAGGAGCAGATGCCCCGACCCCACTTTAGTCTCAGAGACTTGACCTTGA 26688
QY 2515 GCTGTTCTTGAATGACCCAGATGATGATGCTCTGTTTCTCAGGCCAAGGAGCTTGA 2574
DB 26689 GCTGTTCTTGAATGACCCAGATGATGATGCTCTGTTTCTCAGGCCAAGGAGCTTGA 26748
QY 2575 ATGTGCTGACTTATTTATTTTGTGATGATCTGCTGTTTGTGTTTGTGTTTGT 2634
DB 26749 ATGTGCTGACTTATTTATTTTGTGATGATCTGCTGTTTGTGTTTGTGTTTGT 26808
QY 2635 TGTGTTGTTTGTGTTTGTGATGATTTGCTGCTTTCATATATGTAATGCTGTTCTG 2694
DB 26809 TGTGTTGTTTGTGTTTGTGATGATTTGCTGCTTTCATATATGTAATGCTGTTCTG 26868
QY 2695 GGGAACTCCACTGTGCACTGAAGTTATGTAAGAGAGATTTGGCAATGATGCTCCT 2754
DB 26869 GGGAACTCCACTGTGCACTGAAGTTATGTAAGAGAGATTTGGCAATGATGCTCCT 26928
QY 2755 CTATTCAGAGGGGGGTGGGGGGCTTTTCAATGTAATGTTGAGCACTGCTGATTTGAG 2814
DB 26929 CTATTCAGAGGGGGGTGGGGGGCTTTTCAATGTAATGTTGAGCACTGCTGATTTGAG 26988
QY 2815 TCTTCAGTCCCTTCAACACCCAGGCTGGCCACCTCCTCATCTTTCATCTGTGCCAATA 2874
DB 26989 TCTTCAGTCCCTTCAACACCCAGGCTGGCCACCTCCTCATCTTTCATCTGTGCCAATA 27048
QY 2875 AAAAAAAAAAAAAAAAAAAAAA 2896
DB 27049 AAAAAAAAAAAAAAAAAAAAAA 27070

RESULT 5
ID ABLK14000 standard; cDNA; 3353 BP.
AC ABLK14000;
XX 02-JUL-2002 (first entry)
DE cDNA encoding human protein kinase 3700.
KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;
KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;
KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;
KW cell proliferation disorder; cell differentiation disorder; carcinoma;
KW hematopoietic neoplastic disorder; metastatic disorder; leukaemia;
KW cytostatic; antithrombotic; gene; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 157..2043
FT CDS /*tag= a
FT /product= "Protein kinase 3700"
FT /note= "The coding region (not including the terminator
FT codon) is specifically claimed in claims 7 and 28"
XX
XX WO200224921-A2.
XX
XX 28-MAR-2002.
XX
XX 25-SEP-2001; 2001WO-US030115.
XX
XX 25-SEP-2000; 2000US-0234922P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Curtis RAJ, Galvin KM;
XX
XX WPI; 2002-352007/38.

DR P-PSDB: AAU79652.
XX
XX Use of modulators of activity of 3700 protein for making medicament for
PT e.g., modulating protein phosphorylation or cell signaling, or for
PT treating or preventing cellular proliferative and/or differentiative
PT disorders.
XX
XX Claim 28; Fig 1; 115pp; English.
XX
XX The present invention relates to the isolation of a novel human protein
CC kinase designated 3700, and the polynucleotide sequence encoding it. The
CC invention also describes the use of a modulator of the activity of
CC protein kinase (PK) 3700 for making a medicament or pharmaceutical
CC composition for modulating the ability of a cell to phosphorylate an
CC amino acid residue of a substrate protein. Modulators of protein kinase
CC 3700 activity are useful for modulating protein phosphorylation, cell
CC signalling, tumorigenesis, mitogenesis, transcription of a gene,
CC angiogenesis, tissue repair, tissue regeneration, establishment or
CC progression of atherosclerosis, and signalling across the blood-brain
CC barrier. The polynucleotide and polypeptide molecules for protein kinase
CC 3700 may be used as diagnostic targets and therapeutic agents for
CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or
CC curing PK-related disorders and cellular proliferative and/or
CC differentiative disorders (e.g. hematopoietic neoplastic disorders,
CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700
CC polynucleotide sequence can be used to express protein kinase 3700, to
CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for
CC tissue typing, in forensic biology, and as surrogate markers. The present
CC sequence encodes human protein kinase 3700
XX
SQ Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;
Query Match 47.5%; Score 1379; DB 6; Length 3353;
Best Local Similarity 81.8%; Pred. No. 9.3e-295;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;
QY 82 CTACTGCTGCCCAACCACTTCACCTGCGGCTCCCGCACATGAGATCGGTGCTTAC 141
DB 116 CTACTGATTCCTCCCTGCGCCCTTGCTCACCCTCGCTCGCATGAGATCGGTGTTTGG 175
QY 142 TCCAGCGCCGAGCGAGCACTCCCTCGGCTCCCGCCCTGGCTGGAGAAGCGCCGCGCG 201
DB 176 CGCGGCGCTCGGCGCCCACTCCCTCGGCGCGAGCTA-----GCCGCGCGC 223
QY 202 TGGCGAGCGGCTCATCAAGTGGCTTAACTCTGATGAGAAGACAGCGGTGAAGCGCG 261
DB 224 TGGGGAAGGGCTGATCAATCGGCCAAGGCCCTTAAGGAAGACAGCGGTGAAGCGCG 283
QY 262 ACCATCAAAACACACACTCGGCGACCGCTACGAGTTCTTGAGACGCTGGGCAAGGCA 321
DB 284 ACCACCAAAAGCAACACCTCGGCGACCGCTACGAGTTCTTGAGACGCTGGGCAAGGCA 343
QY 322 CCTACGGGAAGGTGAAGAGCAAGAGAGCTCGGGGCGCTGTGTGCCATCAAGTCCA 381
DB 344 CCTACGGGAAGGTGAAGAGAGCTCGGGGCGCTGTGTGCCATCAAGTCCA 403
QY 382 TCAAGGAAGCAAAATCAAAAGATGAGCAGGATGCTGTGACATACGAGGAGATTGAGA 441
DB 404 TCGGGAAGCAAAATCAAAAGATGAGCAGGATGCTGTGACATACGAGGAGATTGAGA 463
QY 442 TCATGCTTCACTCAACCACTCCCACTCATTTGCCATTCATGAAGTTTGTGAGATGCA 501
DB 464 TCATGCTATCACTCAACCACTCCCACTCATTTGCCATTCATGAAGTTTGTGAGAACGA 523
QY 502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
DB 524 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 583
QY 562 AGCGGCAAGCGCTGAGTGAAGCGGAGCGGAGGATTTCTTCCGACAGATGCTGCTGCC 621
DB 584 AGCGGCAAGCGCTGAGTGAAGCGGAGGAGGATTTCTTCCGAGATGCTGCTGCC 643
QY 622 TGCACTAATGCAACCAAGGATGCTTCAACCGAGATTCGAAGCTGGAATAACATCTTC 681

Db 644 TGCACTATGTCATCAAGAGAGTTGTCCACGAGATCTCAAGCTGAGAACTCTCT 703
 Qy 682 TAGATGCAATGAGAAATCAAGATGCTGACTTTGGCTCTCTCAACCTGACCAAG 741
 Db 704 TGAATGCAATGAGAAATCAAGATGCTGACTTTGGCTCTCTCAACCTGACCAAG 763
 Qy 742 GCAAGTTCCTCAAGAGTTCTGTGGAGCCCTCTCTAAGCTCTGAGTAAGTCAAG 801
 Db 764 GCAAGTTCCTGACAGATCTGTGGAGCCCTCTAAGCTCTGAGTAAGTCAAG 823
 Qy 802 GGAAGCCCTATGTTGGCCCAAGAGTGAAGCTGTCTGTGGGCTCTCTGATCAATCC 861
 Db 824 GGAAGCCCTACAGAGCCCAAGAGTGAAGCTGTCTGTGGGCTCTCTGATCAATCC 883
 Qy 862 TGGTGCATGAGCAATGAGCTTTGACGGGAGAGATCAATAAACACTGGTGAAGAAATCA 921
 Db 884 TGGTGCATGAGCAATGAGCTTTGATGGAGATGACATTAAGATCTGATGAAGCAATCA 943
 Qy 922 GTAAAGGGGCTTACCTGAGCCGCAAGCCGTCGATGCTGTGGCTGATCCGATGAGC 981
 Db 944 GCAAGGGGCTTACCTGAGAGCACTTAACCTCTGATGCTGTGGCTGATCCGATGAGC 1003
 Qy 982 TGTATATGATGAAATCCCAAGCTGTGGGCACTGAGAGATGATGCAATGATGAGG 1041
 Db 1004 TGTATATGATGAAATCCCAAGCTGTGGGCACTGAGAGATGATGCAATGATGAGG 1063
 Qy 1042 TCAACTGGGGGTTTACCAAGCCGAGATGAGGAGAAAGCCCTGAGAGGAGGAGCACC 1101
 Db 1064 TCAACTGGGGGTTTACCAAGCCGAGATGAGGAGAAAGCCCTGAGAGGAGGAGCACC 1123
 Qy 1102 CTAGTGTGATCTTTGAGCCGGGCTCTCAATGAGGAGTGTGATGCTGCTCTCGGCCCC 1161
 Db 1124 CTGCAATGATCTGCTGCCCCGCTCTCAATGAGTGTGATGCTGCTGCTGCTGCTGCTG 1183
 Qy 1162 TCCTGAGAAATGAGCAAGGTTGTGAGCTTTCTTCAAGCAGCAGTGTGCGGAGAGTGA 1221
 Db 1184 TCCTGAGAAATGAGCAAGGTTGTGAGCTTTCTTCAAGCAGCAGTGTGCGGAGAGTGA 1243
 Qy 1222 GCACTGATCTGAGGCTGAGAGGAGCAATTTCTTGAAGAAATCCCGAAAGAGAAATGA 1281
 Db 1244 GCACTGATCTGAGGCTGAGAGGAGCAATTTCTTGAAGAAATCCCGAAAGAGAAATGA 1303
 Qy 1282 TGGCTCAAAATCTGCAAGTGAATCCGGCTGAGATCACTTTCTGCGCTGAGAGAGCA 1341
 Db 1304 TGGCTCAATCTCTTCAAGTGAATCCGGCTGAGATCACTTTCTGCGCTGAGAGAGCA 1363
 Qy 1342 GCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAGAGTCTTCACTGCTGAGAGAGTAC 1401
 Db 1364 ACCTCAAGCTGCAAGAGGCAATTTCTCAAGAAAGAGTCTTCACTGCTGAGAGAGTAC 1423
 Qy 1402 AGAGAGACCTTCAAGAAATCAAGCCGCTGATTAATCTTCAAGAGGAGCTGCTGCTG 1461
 Db 1424 AGAGAGACCTTCAAGAAATCAAGCCGCTGATTAATCTTCAAGAGGAGCTGCTGCTG 1477
 Qy 1462 TATCCCTGCTCCCAAGAAAGGCAATTTCTTGAAGATCTGCAAGGAGTCTGCTTACT 1521
 Db 1478 ---CCCTGCTCCCAAGAAAGGCAATTTCTTGAAGATCTGCAAGGAGTCTGCTTACT 1544
 Qy 1522 ACTCTCTTCCAGAGCCGAGAGTCTGAGAACTTTAGAGCCAGATGATGTTTGTGA 1581
 Db 1535 ACTCTCTTCCAGAGCCGAGAGTCTGAGAACTTTAGAGCCAGAGGAGTCTGTTTGTGA 1594
 Qy 1582 GTGGGAGACCCGTTGAGAGCAAGATCTCAAGGCTTCAAGGCTCTCTCTCAAGCAAG 1641
 Db 1595 GTGGGAGATCCCAAGAGCAAGAGCTTCAAGGCTCTCTCTCTCATGCAAG 1654
 Qy 1642 GCATTCTCAAACTCAATGAGCAAGTCTTCCGCAAGCTTGAAGAGCACTACCTCTGCA 1701
 Db 1655 GCATCTCTCAAACTCAATGAGCAAGTCTTCCGCAAGCTTGAAGAGCTACCTCTCTGCA 1714
 Qy 1702 CTTTGGCTCTCTGAGCAACTGAGCTTCTCTTCTGAGAGCCGAGCCGAGCCCTT 1761

Db 1715 CTTTGGCTCTCTGAGCAACTGAGCCCACTGAGCCCTGAGCCGAGCAAGCCCT 1774
 Qy 1762 CAGGGGCTGTGAGTGAAGAGCAAGCATCTGTCTCTCCAGATCTTTGACCAATGATTC 1821
 Db 1775 CAGGGGCTGTGAGTGAAGAGCAAGCATCTGTCTCTCCAGATCTTTGACCAAGTTC 1834
 Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACTGAGGGGC 1881
 Db 1835 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTGTGTGAGCAACTGAGGGGC 1894
 Qy 1882 TTGAGCAGCTCTCCCTGAGAG-----CTGAAAGCATGATGAGGAGAAATCT 1929
 Db 1895 TTGAGGAGCCCTCTCAAGAGGAGCCCTGAAAGCTCTGAGGAGCTGAGGAGAAATCT 1954
 Qy 1930 TGGGAGATGAGCTGCTTTCTTCTGACAGATGCAAGAGTGAATGCAAGCTTACAGAG 1989
 Db 1955 TGGGAGAGAGCTGCTTTCTTCTGACAGATGCAAGAGTGAATGCAAGCTTACAGAG 2014
 Qy 1990 CCTTGAAGATCTGCTCAAGAGCTGAGAGAGAGATGATGCTTATGAGGAGTGA 2049
 Db 2015 CACTGAGAGGCTGCTCAAGAGCTGAGAGAGAGATGAGGAGATGCTTACAGAG 2073
 Qy 2050 GCTTGAAGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092
 Db 2074 GCTTGAAGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2116
 RESULT 6
 ADL14160
 ID ADL14160 standard; cDNA; 3353 BP.
 XX
 AC ADL14160;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Novel human gene 3700 cDNA.
 XX
 KW cytostatic; cardiatic; hypotensive; antilanginal; osteopathic;
 KW antihistatic; antirheumatic; neuroprotective; antiinflammatory;
 KW antipariatic; antiaesthetic; cardiovascular; virucide; analgesic; CNS;
 KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;
 KW nephrotoxic; antihypoid; dermatological; immunomodulator;
 KW cell proliferation inhibitor; cell differentiation inhibitor;
 KW brain disorder; platelet disorder; breast disorder; colon disorder;
 KW kidney disorder; renal disorder; lung disorder; ovarian disorder;
 KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;
 KW thyroid disorder; testes disorder; haematopoietic disorder;
 KW pancreatic disorder; skeletal muscle disorder; skin disorder;
 KW inflammatory disorder; cardiovascular disorder;
 KW endothelial cell disorder; liver disorder; viral disease; pain disorder;
 KW metabolic disorder; neurological disorder;
 KW central nervous system disorder; erythroid disorder;
 KW blood vessel disorder; angiogenic disorder; cancer; heart failure;
 KW hypertension; angina; osteoarthritis; rheumatoid arthritis;
 KW multiple sclerosis; Crohn's disease; psoriasis; asthma;
 KW cell proliferation; cell differentiation; cell growth; cell division;
 KW human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2004058355-A1.
 XX
 PD 25-MAR-2004.
 XX
 PF 25-APR-2003; 2003US-00423543.
 XX
 PR 30-SEP-1998; 98US-00163821.
 PR 27-JAN-1999; 99US-0117580P.
 PR 25-MAR-1999; 99US-00276400.
 PR 30-JUL-1999; 99US-00365162.
 PR 09-SEP-1999; 99US-00392189.
 PR 05-OCT-1999; 99US-00412210.

PR 23-NOV-1999; 99US-00448076.
PR 29-FEB-2000; 2000US-0186061P.
PR 28-APR-2000; 2000US-0200688P.
PR 19-MAY-2000; 2000US-0205447P.
PR 30-JUN-2000; 2000US-00608921.
PR 31-JUL-2000; 2000US-0221925P.
PR 25-SEP-2000; 2000US-0234922P.
PR 25-SEP-2000; 2000US-0235035P.
PR 08-NOV-2000; 2000US-0246669P.
PR 09-NOV-2000; 2000US-00711216.
PR 14-NOV-2000; 2000US-0248325P.
PR 15-NOV-2000; 2000US-0248933P.
PR 22-DEC-2000; 2000US-0257511P.
PR 05-JAN-2001; 2001US-0260156P.
PR 28-FEB-2001; 2001US-00797039.
PR 27-APR-2001; 2001US-00845044.
PR 20-JUL-2001; 2001US-00909743.
PR 31-JUL-2001; 2001US-00920346.
PR 13-AUG-2001; 2001US-00928531.
PR 14-AUG-2001; 2001US-00929218.
PR 15-AUG-2001; 2001US-0312539P.
PR 25-SEP-2001; 2001US-00963159.
PR 08-NOV-2001; 2001US-00008016.
PR 13-NOV-2001; 2001US-00012055.
PR 15-NOV-2001; 2001US-00003690.
PR 30-JAN-2002; 2002US-00060763.
PR 25-MAR-2002; 2002US-00105899.
PR 12-APR-2002; 2002US-00121911.
PR 12-AUG-2002; 2002US-00217168.
PR 22-OCT-2002; 2002US-00278036.
PR 02-JAN-2003; 2003US-00336489.
PR 03-JAN-2003; 2003US-00336153.
PA (MILL-) MILLENNIUM PHARM INC.
PI Kapeller-Libermann R, Hunter JF, Meyers RE, Rudolph-Owen LA;
PI Curtis RAJ, Olande PJ, Teal F, Galvin KM, Chun M, Williamson MJ;
PI Siles-Santiago I, Bandaru R;
XX WPI; 2004-268788/25.
XX P-PSDB; ADL14161.
PT New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,
PT 26176, 26343, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593
PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,
PT heart failure and angina.
XX
XX
PS Claim 1; SEQ ID NO 43; 139bp; English.
XX
XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,
CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,
CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising
CC any one of 40 nucleotide sequences (I). The nucleic acid molecules and
CC polypeptides are useful for diagnosing and treating a subject having a
CC disorder, or a subject at risk of developing a disorder, which is
CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,
CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,
CC m1983, 38555 or 593 activity, such as cellular proliferative and/or
CC differentiative disorders, brain disorders, platelet disorders, breast
CC disorders, colon disorders, kidney (renal) disorders, lung disorders,
CC ovarian disorders, prostate disorders, cervical disorders, spleen
CC disorders, thymus disorders, thyroid disorders, testes disorders,
CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,
CC skin (dermal) disorders, disorders associated with bone metabolism,
CC immune, e.g. inflammatory disorders, cardiovascular disorders,
CC endometrial cell disorders, liver disorders, viral diseases, pain
CC disorders, metabolic disorders, neurological or central nervous system
CC disorders, erythroid disorders, blood vessel disorders or angio-genic
CC disorders (all claimed), e.g. cancer, heart failure, hypertension,
CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's
CC disease, psoriasis, or ashma. The nucleic acid molecules and
CC polypeptides are also useful as modulating agents in regulating a variety
CC of cellular processes, e.g. cell proliferation, differentiation, growth and

CC division. This sequence encodes a novel human protein of the invention.
CC Note: The sequences given in the specification are also available in
CC electronic format from
CC ftp.segdata.uspso.gov/sequence.html?DocID=2004058355.
XX
SQ Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;
Query Match 47.5%; Score 1379; DB 12; Length 3353;
Best Local Similarity 81.8%; Pred. No. 9.3e-295;
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;
82 CTAAGTGTGCGCCGACCCACCTCACTGGGGTCCCGGACACATAGAGTGGCCCTTAC 141
Db 116 CTAAGTGTGCGCCGACCCACCTCACTGGGGTCCCGGACACATAGAGTGGCCCTTAC 175
Qy 142 TCCAGCGCCGAGCGACAGCTCCCTCGGCTCCGCGCTGAGAGCGCCCGCGCG 201
Db 176 CGGGGGGCTCGGGGGGCGGACCTCCCTCGGCGGAGAGCTA-----GCCCGGCGCG 223
Qy 202 TGGCGGAGCGGCTCATAGTCCCTTAACCTTGATGAGAGAGCGGCTGAGCGCGC 261
Db 224 TGGCGGAGAGGCTGATCAAGTCCCGCAAGCCCTTAATGAGAGAGCGGCTGAGCGCGC 283
Qy 262 ACCATCAAAACCAACCTGCGGAGCCGCTACAGAGTTCCTGAGAGCGCTGGAGAGGCA 321
Db 284 ACCACCAAGAGCAACCTGCGGAGCCGCTACAGAGTTCCTGAGAGCCCTGGAGAGGCA 343
Qy 322 CTTACGGGAGAGTGAAGAGAGAGAGAGCTCGGGGCGCTGTTGGCCATCAAGTCA 381
Db 344 CTTACGGGAGAGTGAAGAGAGAGAGAGAGAGCTCGGGGCGCTGTTGGCCATCAAGTCA 403
Qy 382 TCAG 441
Db 404 TCCGAG 463
Qy 442 TCATGCTTACCTCAACCAACCCCACTATTCGATCCATGAGAGTGGAGATGCA 501
Db 464 TCATGCTTACCTCAACCAACCCCACTATTCGATCCATGAGAGTGGAGATGCA 523
Qy 502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 524 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
Qy 562 AGCGGCGAGCGGCTGAGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 621
Db 584 AGCGGCGAGCGGCTGAGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 643
Qy 622 TGCACTACCTGCGCCAG 681
Db 644 TGCACTACCTGCGCCAG 703
Qy 682 TAGATGCAATGAG 741
Db 704 TAGATGCAATGAG 763
Qy 742 GCAAGTCTCTCAG 801
Db 764 GCAAGTCTCTCAG 823
Qy 802 GGAAGCCCTTATGAG 861
Db 824 GGAAGCCCTTATGAG 883
Qy 862 TGGGATGAG 921
Db 884 TGGGATGAG 943
Qy 922 GTAACGGGAGTTCAG 981
Db 944 GCAACGGGAGTTCAG 1003
Qy 982 TGTATATGAG 1041

Db 1004 TGTATGATGAAACCCACCCGCGGACCTGTGAGATGTGGCCATGTCGTGGG 1063
 Qy 1042 TCACTGGGGTTACACCAACCGAGTCCGGGAAACAGAAAGCCCTGCTGAGGGTGGCAAC 1101
 Db 1064 TCAACTGGGGTTACGCCACCGAGTGGGAGAGAGAGGCTCCGATGAGGGTGGCAAC 1123
 Qy 1102 CTAGTGTGATCTTGGCCGGGCTTCATGCGGAGCTGGTTACGTGCTCTCGGCCCC 1161
 Db 1124 CTGGCAGTGAATCTGCGCGCTTCATGCTGACTGGCTCGGCGTTCTCCGCCCC 1183
 Qy 1162 TCCCTGAGAAATGAGCCCAAGTGTGAGCTTCTTCAAGCAGACAGTGCCTGGAGAGTGA 1221
 Db 1184 TCCCTGAGAAATGAGGGCCAAAGGTGTGAGCTTCTTCAAGCAGACAGCTGTGTGGGAA 1243
 Qy 1222 GCACTGATACCTGGGCTGAGAGGCGCAATCTCTTAAAGAAATCCCGAAGAGATGCA 1281
 Db 1244 GCACCAACCTCTGGGCTTGGAGGCGCAAGCTTCTCAAGAAATCCCGAAGAGATGCA 1303
 Qy 1282 TGGCTCAAAATCTGCAAGGTGACCCGGCTGAGATACTCTTCTGCGCTTGGCAAGACA 1341
 Db 1304 TGGCCAGTCTCTCAACAGTGAACGGGTGATGACATGCGCCATCGCTTGGCAAGACA 1363
 Qy 1342 GCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAGTCTTCACTCTGCTGAGGGAGTAC 1401
 Db 1364 ACCTCAAGCTGCCAAAGGCAATTTCTCAAGAAAGTGTCTGCTTGGCAAGAGGTTAC 1423
 Qy 1402 AGAGGACCTCAGAACTGACACCGGCTGATACCTCAAGGGGAGCTGTCCCTGTG 1461
 Db 1424 AGAGGACCTCAGAGCTGACCCCAATCCCTGAGGCCAGGGAGGCTGCC----- 1477
 Qy 1462 TATCCCTGCTCCCAAGAAAGGCAATCTTAAAGAGTCTGACAGCGTGAATCTGGTTACT 1521
 Db 1478 ---CCCTGCTCCCAAGAAAGGCAATTTCAAGAAAGCCCGACAGCGAGTCTGGTACT 1534
 Qy 1522 ACTCTCTCCAGAGCCCAAGGCTGTGGGAACTCTTAAAGCAGCAAGTGTGTTGTA 1581
 Db 1535 ACTCTCTCCAGAGCCCAAGTGAATCTGGGAGCTCTTGAAGCAGAGCGACGTGTTGTA 1594
 Qy 1582 GTGGGAGCCCTGGAGAGAGTCTCAGAGGCTTCAAGGGCTCTCTCAACCGCAAG 1641
 Db 1595 GTGGGAGTCCCAAGAGAGAGGCTCCGCAAGCTTCAAGGGCTCTCTCATGCAAG 1654
 Qy 1642 GCATTCTCAAACTCAATGCAAGTCTCCGCAAGCTTGAAGAGCACTACCTTAGCA 1701
 Db 1655 GCATCTCAAACTCAATGCAAGTCTCTCCAGACAGCTTGAAGTCTGCGGCCCAACCA 1714
 Qy 1702 CTTTGGCTCTGAGCAACTGAGCTCTCTCCATCTGACAGCCCGCCAGCCCTT 1761
 Db 1715 CTTTGGCTCTGAGCAACTGAGCTCTCTCCATCTGAGCTCTTGAAGCAGCTGACCTT 1774
 Qy 1762 CAGGGGCTGTGAGTGAAGACAGCATCTGTCTTCCAGATCTTTGAAGCAATTGACCTTGC 1821
 Db 1775 CAGGGGCTGTGAGTGAAGACAGCATCTGTCTTCCATCTTGAAGCAGCTGACCTTGC 1834
 Qy 1822 CTGAAGCTTCTCCGAAACCCCACTGAGGGGCTGTGAGTCTGTGAGCAAACTGAGGGGC 1881
 Db 1835 CTGAAGCTTCTCCGAAAGCCCACTGCGGGCTGTGTGTCTGTGAGCAAACTTCAACGGGC 1894
 Qy 1882 TTGAGCAGCTCTCTCAGAAAG-----GTCTGAAGCAATGTGTGAGAGAACTCTT 1929
 Db 1895 TTGAGAGAGCCCTCTCAGAGAGCCCTGGAAGCTGTGAGGCGGTGGGGGAGAGTCTT 1954
 Qy 1930 TGGGGGATAGCTGCTTTTCTTGAACAGCTGCAAGAGTGAATGCAAGCTTACAGACAG 1989
 Db 1955 TGGGGGACAGCTGCTTTTCTTGAACAGCTGCAAGAGTGAATGCAAGCTTACAGACAG 2014
 Qy 1990 CCTAAGAAATCTGTCAAAAGCTCAGCTGAGAGAGGAGAGATGTGCTGATGTGGGGTAG 2049
 Db 2015 CACTGAGAGTCTGTCAAAAGCTCAGCTGAGAGTGAATGCAATGCTCCAG--CCGGTCA 2073
 Qy 2050 GCTGTGAGAGGTTTGCAGAGAACTCTGGGTGGATTTCTCC 2092
 Db 2074 GCTCTCAGATGCAAGTGTGTGACCCCGAGGGGAGATGCTTTC 2116

RESULT 7
 ID AAK94280
 AAK94280 standard; cDNA; 3395 BP.
 AAK94280;
 06-NOV-2001 (first entry)
 Human full-length cDNA, SEQ ID NO: 2918.
 Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 Homo sapiens.
 EP1130094-A2.
 05-SEP-2001.
 07-JUL-2000; 2000EP-00114089.
 08-JUL-1999; 99JP-00194486.
 11-JAN-2000; 2000JP-00118774.
 02-MAY-2000; 2000JP-00183765.
 (HELI-) HELIX RES INST.
 Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y,
 Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 WPI; 2001-52425/58.
 P-PSDB; AAM93360.
 830 Primers useful for synthesizing full length cDNA clones and their use
 in genetic manipulation.
 Claim 8; SEQ ID NO 2918; 1380bp + Sequence Listing; English.
 The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been isolated
 and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 been determined. Primers for synthesizing the full length cDNA are useful
 for clarifying the function of the protein encoded by the cDNA. The full
 length clones were obtained by construction of full length enriched cDNA
 libraries that were synthesised by the oligo-capping method. The primers
 enable the production of the full length cDNA easily without any special
 methods. The present sequence is a full length human cDNA of the
 invention. Note: The sequence data for this patent did not form part of
 the printed specification, but was obtained in CD-ROM format directly
 from BPO
 Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
 Query Match 47.4%; Score 1375.8; DB 4; Length 3395;
 Best Local Similarity 81.7%; Pred. No. 476-294;
 Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
 82 CTACTGTGCGCGACCACTTCACCTCGCGGTCCCGGACCATGAGTGTGAGCTTAC 141
 Db 77 CTACTGATTCCTCCGCGGCTTGTCTCACTCTGCTGCGCATGAGTGTGCTGTTTCG 136
 Qy 142 TCCAGCGCCGAGGCAAGCTTCCCTGCGCTTCCGCGCTGAGCTCGGAGAGCGCCGCGC 201
 Db 137 CGCGGCTCTCGGCGCCCACTCCCTCGGCGCGAGGCTA-----GCCGCGCGC 184
 Qy 202 TGGGGAAGGGCTCATCAAGTGTGCTTAACTCTGATGAAGAGACAGCGGTGAAGCGGC 261
 Db 185 TGGGGAAGGGCTCATCAAGTGTGCTTAACTCTGATGAAGAGACAGCGGTGAAGCGGC 244
 Qy 262 ACCATCAAAACACAACTTGGGCAACCGCTACGAGTTCCTGAGAGCGTGGGCAAGGCA 321
 Db 245 ACCACCAAGAGCAAACTTGGGCAACCGCTACGAGTTCCTGAGAGCGTGGGCAAGGCA 304

QY	322	CTCAGCGGAAGGTGAAGAAAGGACAGAGAGAGTCCGGGGGCTCTGGTGGCCATCAAGTCA	381
Dp	305	CTTACGGGAAGGTGAAGAAAGGCGGGAGAGCTCCGGGGCGCTGGTGGCCATCAAGTCA	364
QY	382	TCAGAAAGACAAAATCAAAAGTAGAGCAGAGTCTGCTGCATACAGAGGAGATTGAGA	441
Dp	365	TCGGGAAGAGCAAAATCAAAAGTAGAGCAAGATCTGATGCACTTACGGAGGGAGATTGAGA	424
QY	442	TCATGTCCTCACTCAACACACCCCACTCATTTGGCATTCATGAACTGATTGGAATPACA	501
Dp	425	TCATGTCACTCAACACACCTCACTCATATTGGCATTCATGAACTGATTGGAACACGA	484
QY	502	GCAAGATTGTGATTTGTCAATGAGATATGCGCAGCGGAGATCTGATGATTACATCAATG	561
Dp	485	GCAAGATGCTGATGCTCATGAGGATATGCGCAGCGGAGACCTTTATGATCATACAGCG	544
QY	562	AGCGGCCACGGCTGATGATGAGCGGAGCGCCAGGCATTCTTCCGACAGATCGTGTCTGCC	621
Dp	545	AGCGGCACACCTCAAGTAGAGCGCGAAGCTAAGGCATTCTTCCGAGAGATCGTGTCTGCC	604
QY	622	TGCATCTACTGCGACACAGAACGGGATCGTTACACGAGATCTCAAGCTGAGAAAACATCCCTT	681
Dp	605	TGCATATTGCGCATATGAAACAGAGTTGTTCACCGAGATCTCAAGCTGAGAACATCTCTT	664
QY	682	TAGATGCCAATGAGAAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACACAAAG	741
Dp	665	TGATATGCCAAATGAGAAATCAAGATTGCTGACTTGGCCTCTCCAACTGTACATCAAG	724
QY	742	GCAAGTTCCTCOAGACGTTCTGTGGGAAACCCCTCTAGCCCTGCGCTGAGATATGATCACG	801
Dp	725	GCAAGTTCCTGAGACATTTCTGTGGGAAACCCCTCTTAGTCCCTGCGCAAGATTGTCAATG	784
QY	802	GGAAGCCCTATATGTGGGCCCAAGAGTGAGCAGCTGTCTCTGGGCCCTTCTCCTGTACATCC	861
Dp	785	GGAAGCCCTTACACAGGCCCAAGAGTGAGCAGCTGTCCCTGGGTCTTCTCCTGTACATCC	844
QY	862	TGTGTGCATGSCAACCATGCCCCCTTTTGAACGGGCAAGATCATPAAAACATGSGTGAACAAATCA	921
Dp	845	TGTGTGCATGAGCAACCATGCCCCCTTTTGAATGGGCAATGACATPAAATCTTATGTGAACACAGATCA	904
QY	922	GTAACGGGGCTTACCGGTAGCGCGCCCAAGCCGTCGATGCGCTGAGGCGCTGATCCGGTGGC	981
Dp	905	GCAACGGGGCTTACCGGAGGACCACTTAACCTCTGATGCTGTGGCTGATGCCCTGATCCGGTGGC	964
QY	982	TGTTATATGTGAACCCCAACCCCTCGGGGCACACTGAGAGATGTAGCCAGTCAATGGTGGG	1041
Dp	965	TGTTATATGTGAACCCCAACCGCGGGGCACCTGTGAGATGTAGCCAGTCACTGGTGGG	1022
QY	1042	TCACATCGGGGTATACACCAACCGGAGTCCGGGGAAACAGAAAGCCCTGCTGTAGGGTGGGACCC	1101
Dp	1025	TCACATCGGGGTATACACCAACCGGAGTGGGAGAGACAGAAAGCTTCGATTAAGGATGGGACCC	1084
QY	1102	CTAGTGTGATCTTTGGCCGGGCTCTCATATGCGGAGACTGTGTACGTGACTCTGCGGCCCCC	1161
Dp	1085	CTGGCAGTGAATCTGTGCCCGGGCTCTCAATGGCTGACTGGCTCCGGGGTTCCTCCGGCCCCC	1144
QY	1162	TCTCTGAGAAATGAGACCCCAAGGTGTGACGCTTCTTCAAGCAGCAGTGTCCGGAGGTGGAA	1221
Dp	1145	TCTCTGAGAAATGAGAGCCCAAGGTGTGACGCTTCTTCAACACACATGCACTGTGTGGGAA	1204
QY	1222	GCACTGTACCTCGGGGTGGAGCGGCAACATTCCTTTAAGAAAGTCCCGAAGAGAAATGACA	1281
Dp	1205	GCACTACCCCTTGGGCTGGAGCGGCAACATTCGCTTCAAGAAAGTCCCGCAAGAGAAATGACA	1264
QY	1282	TGGCTCAAAATCTGCAAGGTGACCCCGCTGAGATACCTCTTCTTCCGCTTGGCAAGACA	1341
Dp	1265	TGGCCCAAGTCTCTCCACAGTGACACCGCTGATGACATGCGCOATGCGCTTGGCAAGACA	1324
QY	1342	GCCTTAAGCTTCCGAAAGGCACTTCTCAAGAAAAAGTCTCTTACTCTGTCAAGGAGGTAC	1401
Dp	1325	ACCTCAAGCTGCCAAAGGGCAATTTCTCAAGAAAGAGTCAAGCTCTGTCAAAAGGGGTAC	1384

QY	1402	AGAGAGACCCCTCAGAACTCAGACCGGTGCTCTGATCTTCCAGGGGACGCTGTCCCTGCTG	1461
Db	1385	AGAGAGACCCCTCCGGAAGCTCAGCCCAATCCCTGGAGACCCAGGGACAGGCTGCC-----	1438
QY	1462	TATCCCTGCTCCAGGAAAGGCATCTCTTAAGAGTCTCGACAGCTGAAATCTGTGTTACT	1521
Db	1439	---CCCTGCTCCCAAGAGGCAATTTCTCAAGAGCCCGCAGCGAGCTGTGGTACT	1495
QY	1522	ACTCTCTCCAGACCCAGGAGCTGTGGGGAACTTTAGACGCCAGTAGTAGTGTGGTGA	1581
Db	1496	ACTCTCTCCAGACCCAGGAAATCTGTGGAGCTCTTGGAGCAGGGGCAAGTGTGGTGA	1555
QY	1582	GTGGGGACCCCGTGGAGCAGAGTCTCAACAGCTTACAGGGCTCTCTCCACCGCAAG	1641
Db	1556	GTGGGGATCCCAAGAGACAGAGGCTCCGCAAGCTTACAGGGCTCTCTCATATGCAAG	1615
QY	1642	GCATTTCTCAACTCAATGCGAAGTTCTCCGCAACGCTTACAGAGGACCTACCCCTTACA	1701
Db	1616	GCATCTCAAACTCAATGGCAAGTTCTCCAGACAGCTTGAAGCTTGAGCTGCGGCCCA	1675
QY	1702	CCTTTGGCTCCCTGGACCAACTGGCTCTCTCCATCTGTGACAGCCCGGCCAGCCCT	1761
Db	1676	CCTTGGCTCCCTGGATGAACCTGCCCCACCTGCCCCCTGGCCCCGGGACAGCGACCT	1735
QY	1762	CAGGGGCTGTAGTAGAGACAGCATCTCTGCTCCAGTCCCTTGAACAAATTGACTTGC	1821
Db	1736	CAGGGGCTGTAGTAGAGACAGCATCTCTGCTCTTGAAGTCTTTGACCAAGCTGACTTGC	1795
QY	1822	CTGAACGTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGACAACTCTGAGGGGC	1881
Db	1796	CTGAACGGCTTCCAGAACCCCACTGAGGGGCTGTGTCTGTGACAACTCTGAGGGGC	1855
QY	1882	TTGAGCAGCTCTCTCTCAGAA-----GTCTGAAGCGATGGTGGCAGGAATCTT	1929
Db	1856	TTGAGGGGCCCCCTCAGAGGGCCCTGGAAGCTGCTGAGAGCGCTGCGCGCAGATCTT	1915
QY	1930	TGGGGGATGATGCTTTTCTGTGACAGACTGCCAAGAGGTGATGTCAGACCTACAGACAAG	1989
Db	1916	TGGGGGATGATGCTTTTCTGTGACAGACTGCCAAGAGGTGATGTCAGACCTACAGACAAG	1975
QY	1990	CCCTAGGAATCTGCTCAAACTCAGCTGAGAGAAAGGAGATGTGTCCCTAGTATGGGGTAG	2049
Db	1976	CACCTAGAGGCTGTGCTCAAACTCAGCTGAGTGAATGAGATGTGCCAG-----CCGGTTCAG	2034
QY	2050	GCTCTGAGAGGTTTTCAGAGAACTCTGAGTGGATCTCTCC	2092
Db	2035	GCTCTCAGATGACAGCTGTTCACCCCGAGGGGAGATGCTTC	2077
RESULT 8			
ADL30885	ADL30885 standard; cDNA; 3395 BP.		
XX	ADL30885;		
AC	ADL30885;		
XX	20-MAY-2004 (first entry)		
DT	XX		
XX	Full length human cDNA clone Segid 2918.		
DE	XX		
KW	human; medicine; signal transduction; glycoprotein; transcription;		
KM	oligo-capping method; ss; gene.		
XX	Homo sapiens.		
OS	XX		
PN	EP1396543-A2.		
XX	10-MAR-2004.		
PD	XX		
XX	07-JUL-2000; 2003EP-00025638.		
PF	XX		
PR	08-JUL-1999; 99JP-00194486.		
PR	11-JAN-2000; 2000JP-00118774.		
RR	02-MAY-2000; 2000JP-00183865.		

PR 07-JUL-2000; 2000EP-00114089.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
DR P-PSDB; ADL30886.
XX
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
length human cDNAs.
PS Example 1; SEQ ID NO 2918; 1340bp; English.
XX
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
XX
SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
Query Match 47.4%; Score 1375.8; DB 12; Length 3395;
Best Local Similarity 81.7%; Pred. No. 4.7e-294;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
QY 82 CTAAGTGTGCGCCGACCTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 141
DB 77 CTAAGTGTGCGCCGACCTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 136
QY 142 TCCAGCGCGCCGACCTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 201
DB 137 CGGCGCGCGCCGACCTTCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTAC 184
QY 202 TGGCGGCGCGGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 261
DB 185 TGGCGGCGGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 244
QY 262 ACCATCAACAACAACCTGCGGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 321
DB 245 ACCATCAACAACAACCTGCGGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 304
QY 322 CTTACGGGAAAGTGAAGAGGACGAGAGGCTCGGCGGCTGAGTGGTGGTGGTGGTGG 381
DB 305 CTTACGGGAAAGTGAAGAGGACGAGAGGCTCGGCGGCTGAGTGGTGGTGGTGGTGG 364
QY 382 TCCGAAAGAGCAAAATTAAGATGAGAGAGTGTGTCGACATGCGAGGAGATTGAGA 441
DB 365 TCCGAAAGAGCAAAATTAAGATGAGAGAGTGTGTCGACATGCGAGGAGATTGAGA 424
QY 442 TCAATGTTCACTCAACACCGCCGACATGTCGATGCAATGAGTGGTGGTGGTGGTGG 501
DB 425 TCAATGTTCACTCAACACCGCCGACATGTCGATGCAATGAGTGGTGGTGGTGGTGG 484
QY 502 GCAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 485 GCAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
QY 562 AGGCGGCAAGGCTGAGTGAAGGCGGAGGCTGATTTCTTCCGACATGTCGTCGCG 621
DB 545 AGGCGGCAAGGCTGAGTGAAGGCGGAGGCTGATTTCTTCCGACATGTCGTCGCG 604
QY 622 TGCACTAGTGCACCAAGAGGAGTGTTCACGAGATCTCAAGTGGTGGTGGTGGTGGTGG 681
DB 605 TGCACTAGTGCACCAAGAGTGTTCACGAGATCTCAAGTGGTGGTGGTGGTGGTGGTGG 664
QY 682 TAGATGCAATGGAACATCAAGATTTGATGATTTGATGATTTGATGATTTGATGATTTG 741

DB 665 TGATGCGCAATGGGAATATCAAGATTTGATGATTTGATGATTTGATGATTTGATGATTTG 724
QY 742 GCAAGTTCCTCAGACCTTGTGTGGAGCCCTCTCTCAAGCTCGCTGAGATAGTCAAG 801
DB 725 GCAAGTTCCTCAGACCTTGTGTGGAGCCCTCTCTCAAGCTCGCTGAGATAGTCAAG 784
QY 802 GGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 861
DB 785 GGAAGCCCTATGTGGGCGCCAGAGGTGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 844
QY 862 TGGTGCATGACCATCCTGCTTGAAGGGAGATGATTAAGACCTGTGAGGAATATCA 921
DB 845 TGGTGCATGACCATCCTGCTTGAAGGGAGATGATTAAGACCTGTGAGGAATATCA 904
QY 922 GTAAGCGGCTTACCTGAGACCGCCCAAGCGCTGAGTGGTGGTGGTGGTGGTGGTGGTGG 981
DB 905 GCAAGCGGCTTACCTGAGACCGCCCAAGCGCTGAGTGGTGGTGGTGGTGGTGGTGGTGG 964
QY 982 TGTATATGTGAACCCCAAGCGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1041
DB 965 TGTATATGTGAACCCCAAGCGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1024
QY 1042 TCAACTGAGGTTACACACCGGAGTGGGAGGACAGAAAGCTGAGGAGTGGGAGC 1101
DB 1025 TCAACTGAGGTTACACACCGGAGTGGGAGGACAGAAAGCTGAGGAGTGGGAGC 1084
QY 1102 CTAAGTGTGACTTTGCGCGGCGCTCATGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1161
DB 1085 CTGAGATGACTTGTGCGCGGCGCTCATGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1144
QY 1162 TCTGAGAAATGAGACCAAGGTGTGAGCTTCTTCAAGCAGACGTCGCGGAGGTGAA 1221
DB 1145 TCTGAGAAATGAGACCAAGGTGTGAGCTTCTTCAAGCAGACGTCGCGGAGGTGAA 1204
QY 1222 GCACTGTACTGTGGCTGAGAGCGGACATTTCTTGAAGAGTCCCGAAAGAGATGACA 1281
DB 1205 GCACTGTACTGTGGCTGAGAGCGGACATTTCTTGAAGAGTCCCGAAAGAGATGACA 1264
QY 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGTGAACCTTCTTGGCTGCGGCAAGAGA 1341
DB 1265 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGTGAACCTTCTTGGCTGCGGCAAGAGA 1324
QY 1342 GCTTTAAGCTTCCGAAAGGACATTTCTCAAGAAAGTCTTCACTGCTGAGGAGGTAC 1401
DB 1325 ACCTCAAGCTTCCGAAAGGACATTTCTCAAGAAAGTCTTCACTGCTGAGGAGGTAC 1384
QY 1402 AGAGGACCTTCAAGAACTCAAGACGCTGTGATCTTCAAGGACGCTGTGCTGCTG 1461
DB 1385 AGAGGACCTTCAAGAACTCAAGACGCTGTGATCTTCAAGGACGCTGTGCTGCTG 1438
QY 1462 TATCCCTGCTCCCAAGGAAAGGACATCTTGAAGAGTCTGCAAGCGTGAATCTGATTACT 1521
DB 1439 TATCCCTGCTCCCAAGGAAAGGACATCTTGAAGAGTCTGCAAGCGTGAATCTGATTACT 1495
QY 1522 ACTCTCTCCAGAGCCGAGGAGTCTGAGGAACTTTAGACGCGCAGATGATGTTTGTGA 1581
DB 1496 ACTCTCTCCAGAGCCGAGGAGTCTGAGGAACTTTAGACGCGCAGATGATGTTTGTGA 1555
QY 1582 GTGGGAGACCCGTGAGCAAGAGTCTTCAAGGCTTCAAGGCTTCTTCTTCAAGGAG 1641
DB 1566 GTGGGAGATCCCAAGGAGCAAGAGCTTCAAGGCTTCAAGGCTTCTTCTTCAAGGAG 1615
QY 1642 GCAATTCGAAATCAATGAGGAGTGTCCGACAGGCTTGAAGAGCACTAACCTTGA 1701
DB 1616 GCAATTCGAAATCAATGAGGAGTGTCCGACAGGCTTGAAGAGCACTAACCTTGA 1675
QY 1702 CTTTGGCTCCCTGAGCAACCTGAGCTTCTTCAAGCTTCAAGGCTTCTTCTTCAAGGAG 1761
DB 1676 CTTTGGCTCCCTGAGCAACCTGAGCTTCTTCAAGCTTCAAGGCTTCTTCTTCAAGGAG 1735
QY 1762 CAGGGGCTGTGAGTGAAGAGAGATCTGTGCTTCTTCAAGCTTCTTGAAGCAATTTGAG 1821
DB 1736 CAGGGGCTGTGAGTGAAGAGAGATCTGTGCTTCTTGAAGCAATTTGAG 1795

QY 1822 CTGAACCTCTTCCGAAACCCCACTGAGGGCTGTGTCTGTGACAACTGAGGGGC 1881
 DB 1796 CTGAACGGCTTCCAGAGCCCTCACTGCGGGCTGTGTCTGTGACAACTGAGGGGC 1855
 QY 1882 TTGAGCAGCCTCCCTCAGAAAG-----GCTGAAGCGATGGGAGGAATCCT 1929
 DB 1856 TTGAGAGGCCCCCTCAGAGGGCCCTGGAAGCTGCTGAGGCGCTGGCGGAGGAATCCTT 1915
 QY 1930 TGGGGGATGAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGCACGCTTACAGACAG 1989
 DB 1916 TGGGGGACAGCTGCTTTTCCCTGACAGACTGCCAAGAGTGCACGCTTACAGACAG 1975
 QY 1990 CCTTGAAGATCTGTCAAGCTCAGCTGAGAGAGGAGATGTGCTTATGTGGGTAG 2049
 DB 1976 CACTGAGGGTCTGCTCAAGAGCTCAGAGTGGGAGGATGCTCCAG-CCCGATCAG 2034
 QY 2050 GCTCTGAGAGGGTTTTCAGAGGAACCCCTGGTGGATTCCTCC 2092
 DB 2035 GCTCTCAGATGCACTGCTGTGCACTCCGAGGGAGATGCTTTC 2077

RESULT 9
 ABD33085
 ID ABD33085 standard; cDNA; 3404 BP.
 AC ABD33085;
 XX
 XX 18-NOV-2004 (first entry)
 DE Human cancer-associated (CA) cDNA HR07-004.
 XX
 XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW B6; cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 XX MO2004058146-A2.
 PN
 XX 15-JUL-2004.
 PD
 XX 15-DEC-2003; 2003WO-US040081.
 PF
 XX 17-DEC-2002; 2002US-00322281.
 PR
 XX (SAGR-) SAGRES DISCOVERY INC.
 PA
 PI Morris DM, Malandro MS;
 XX
 XX WPI; 2004-499109/47.
 DR P-PSDB; AB084757.
 DR
 XX Novel human cancer associated protein encoded within open reading frame
 PT of cancer associated gene, useful as targets for diagnosing cancer.
 XX
 PS Claim 1; SEQ ID NO 25; 1822bp; English.
 XX
 XX The invention relates to cancer-associated proteins (CAP) and the cancer-
 CC associated (CA) nucleic acids encoding them. The invention also relates
 CC to a method for treating cancers involving administering to a patient an
 CC inhibitor of CAP, and a method of screening for anticancer activity in a
 CC potential drug involving providing a cell that expresses a CA gene,
 CC contacting a tissue sample derived from a cancer cell with an anticancer
 CC drug candidate and monitoring the effect of the anticancer drug candidate
 CC on expression of the CA gene. The CAP proteins are useful for detecting
 CC cancer associated with expression of a CAP protein in a test cell sample
 CC and for screening for a bioactive agent capable of modulating the
 CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 CC cancer, involving determining the expression of a CA nucleic acid in a
 CC tissue. This sequence represents human CA cDNA of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3404 BP; 723 A; 1044 C; 895 G; 742 T; 0 U; 0 Other;
 Query Match 47.4%; Score 1375.8; DB 13; Length 3404;
 Best Local Similarity 81.7%; Pred. No. 4.7e-294;
 Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;

QY 82 CTACTGCTGCCCAACCACTCACTGCGGCTCCCGACATGAGTGGGCTTAC 141
 DB CTACTGATTTCCCTGCGGCTTGTCTCACTCTGCTGCGCATGAGATCGCTGTTTTCG 146
 QY 142 TCAGAGCGCCGACCGACGCGCTCCCGGCTCCCGCTCGAGAGGCGCGCGCG 201
 DB CGGCGGCTCCGCGCCCACTCCCTCGGCGGAGCTA-----GCCCGGCGC 194
 QY 202 TGGGGAAGGGCTATCAATGCTTAACTCTGATGAGAGAGAGAGAGAGAGAGAG 261
 DB TGGGGAAGGGCTATCAATGCTTAACTCTGATGAGAGAGAGAGAGAGAGAGAGAG 254
 QY 195 TGGGGAAGGGCTATCAATGCTTAACTCTGATGAGAGAGAGAGAGAGAGAGAG 254
 DB ACCACCAAGAGCAACCTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAGAGAG 314
 QY 262 ACCATCAAGAGCAACCTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAGAG 321
 DB ACCACCAAGAGCAACCTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAGAGAG 314
 QY 322 CCTACGGGAAG 381
 DB CCTACGGGAAG 374
 QY 315 CCTACGGGAAG 374
 DB 382 TCAGAAAGAGCAAAATCAAGATGAGAGAGAGATGCTGACATACGAGAGAGATTGAGA 441
 DB 375 TCCGGAAGAGCAAAATCAAGATGAGAGAGAGATGATGACATACGAGAGAGATTGAGA 434
 QY 442 TCATGCTTCACTCAACCACTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAG 501
 DB TCATGCTTCACTCAACCACTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAGAG 494
 QY 435 TCATGCTTCACTCAACCACTGCGGACCGCTCAAGATTCCTGAGAGAGAGAGAG 494
 DB 502 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
 DB 495 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
 QY 562 AGCGGCAAGGCTGAGTGAAGCGGAGCGGAGGATTTCTTCCAGAGATGATGATGATGAT 621
 DB 555 AGCGGCAAGGCTGAGTGAAGCGGAGCGGAGGATTTCTTCCAGAGATGATGATGATGAT 614
 QY 622 TGCACTACTGCAACCAAGCGGATGCTTCAAGATTCCTGAGAGAGAGAGAGAGAG 681
 DB 615 TGCACTACTGCAACCAAGCGGATGCTTCAAGATTCCTGAGAGAGAGAGAGAGAG 674
 QY 682 TAGATGCAATGGAATCATCAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAG 741
 DB 675 TAGATGCAATGGAATCATCAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAG 734
 QY 742 GCAAGTTCCTCAAGAGCTTCTGTGAGAGCGCTCTTCAAGCGCTGAGATGATGATGAT 801
 DB 735 GCAAGTTCCTCAAGAGCTTCTGTGAGAGCGCTCTTCAAGCGCTGAGATGATGATGAT 794
 QY 802 GGAAGCCCTATGATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
 DB 795 GGAAGCCCTATGATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
 QY 862 TGGTGCATGAGCAGATGCTTGTGAGAGGAGAGATCATTAACATGATGATGATGATGAT 921
 DB 855 TGGTGCATGAGCAGATGCTTGTGAGAGGAGAGATCATTAACATGATGATGATGATGAT 914
 QY 922 GTTAAAGGGGCTTACCGTGAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
 DB 915 GCAAGCGGCTTACCGTGAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 974
 QY 982 TGTAAATGATGAACCCACCGCTGAGGAGCACTGAGAGATGATGATGATGATGATGATGAT 1041
 DB 975 TGTAAATGATGAACCCACCGCTGAGGAGCACTGAGAGATGATGATGATGATGATGATGAT 1034
 QY 1042 TCAACTGAGGTTTACCAACCGAGAGTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101

Db 1035 TCACTGGGGGCTACCCAGCCGAGTGGAGACAGAGGCTCCGCGATGAGGCTGGGACAC 1094
Qy 1102 CTAGTGGTGACTTTGGCCGGGCTCTCATGGCGGACTGGTTACGTGCTCTCGCGCCCC 1161
Db 1095 CTGGCAGTGAATCTGGCCGGGCTCCAGTGAATGGCTCGGGGTTCTCCGCCCCCC 1154
Qy 1162 TCCTGAGAGATGAGCCAGAGTGTGACGCTTCTTCAAGCAGACGTCGCGGAGGTGAA 1221
Db 1155 TCCTGAGAGATGAGGCGCAGAGTGTGACGCTTCTTCAAGCAGACGTCGTCGGGGGAA 1214
Qy 1222 GCACTGATACCTGGGCTGGAGGGGACATTTCTTAAAGATCCCGAAGAGATGACA 1281
Db 1215 GCACACCCCTGGGCTGGAGGGCAGCATTCCTCAAGAAATCCCGAAGAGATGACA 1274
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCCGGCTGAGATACCTCTTCTCGCCCTGGCAAGACA 1341
Db 1275 TGGCCAGTCTCTCAAGTGAACCGCTGATGACATGCGCCATGCGCCCTGGCAAGACA 1334
Qy 1342 GCCTTAAGCTTCCGAAAGGCAATTCCTCAAGAAAGTCTCTCACTCGTCAGGGAGGTAC 1401
Db 1335 ACCTCAAGCTCCAAAGGGCATTCCTCAAGAAAGTGTGACGCTCTGCAAGAGGGGTAC 1394
Qy 1402 AGAGGACCTCTCAGAACTCAGACCGGTGCTGATATCTCAGGGGACGCTCTGCTG 1461
Db 1395 AGAGGACCTCTCAGAGCTCAGCCCAATCCTTGGAGGCCAGGGGAGGCTGCC----- 1448
Qy 1462 TATCCCTGCTCCGAAAGGAGGCAATCTTAAAGTCTGACAGCGTGAATCTGCTTACT 1521
Db 1449 ---CCCTGCTCCGAAAGGAGGCAATTCCTCAAGAAAGCTCCGACAGCGAGTGTGCTACT 1505
Qy 1522 ACTCCTCTCAGAGCCGAGGAGTCTGGGGAATCTTGAAGCGCAGTGAATGTTTGA 1581
Db 1506 ACTCCTCTCCGAGCCGAGGATCTGGGAGCTCTTGAAGCGCAGGCGACGTGTTTGA 1565
Qy 1582 GTGGGACCCCGTGGAGAGAGTGTCTCAGAGCTTCAAGGCTCTCTTCAAGCGAAG 1641
Db 1566 GTGGGAGTCCCAAGAGAGAGGCTCCGAGGCTTCAAGGCTCTCTTCAAGCGAAG 1625
Qy 1642 GCATTCTCAACTGAATGGAGAGTTCCTCCGACAGCGTGAAGGCACTACCCCTGACA 1701
Db 1626 GCATCTCAACTGAATGGAGAGTTCCTCCGACAGCGTGAAGTTCGCGGCCCAACCA 1685
Qy 1702 CCTTGGCTCCCTGAGCAACATGGGCTCTCTCCATCTCGAGCCGCGGCCAGCCCT 1761
Db 1686 CCTTGGCTCCCTGAGTGAATCGGCCCACTCCGCCCTCGCGGCCAGCGGACCT 1745
Qy 1762 CAGGGGCTGTGAGTGAAGAGAGCATCTGTCTTCGAGTCTTGAACCATTTGACTTGC 1821
Db 1746 CAGGGGCTGTGAGGAGAGAGCATCTGTCTTCGAGTCTTGAACCATTTGACTTGC 1805
Qy 1822 CTGAAGCTTTTCCGAAACCCTGAGAGGCTGTGTCTGTGAGCAACCTGAGGGGC 1881
Db 1806 CTGAAGGCTTCCAGAGCCCTGAGGCTGTGTCTGTGAGCAACCTGAGGGGC 1865
Qy 1882 TTGAGCAGGCTCCCTCAGAG-----GTTGAGAGAGTGTGAGGAGAGATCT 1929
Db 1866 TTGAGAGGCCCCCTCAGAGGCCCCCTGAGAGCTGTGAGGCGCTGAGGAGAGATCTT 1925
Qy 1930 TGGGGAATAGCTGTTTTTCTTGAACAATGCAAGAGTGAATGCAAGCTTACAGACAG 1989
Db 1926 TGGGGAACAGCTGTTTTTCTTGAACAATGCAAGAGTGAATGCAAGCTTACAGACAG 1985
Qy 1990 CCCGAGAAATGTGTTCAAAAGTCAAGTGAAGAGAGAGATGTGCTTGAATGGGGTNG 2049
Db 1986 CACTGAGAGTGTGTTCAAAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2044
Qy 2050 GCTTGAAGAGGTTTGAAGAGAACTGTGGTGGATTCCTCC 2092
Db 2045 GCTTGAAGTGAAGTGTGCAACCCGAGGGAGATGCTTTC 2087

RESULT 10
ABX71420

ID ABX71420 standard; cDNA; 3443 BP.
XX
AC ABX71420;
XX
DT 14-APR-2003 (first entry)
XX
DE Human cell cycle-associated cDNA from clone DKFZp473c3_713.
XX
KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
OS Homo sapiens.
XX
PN W0200112659-A2.
XX
PD 22-FEB-2001.
PE 18-AUG-2000; 2000MO-IB001496.
PR 18-AUG-1999; 99US-014949P.
PR 28-SEP-1999; 99US-0156503P.
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
PI Wiemann S;
PI
DR WPI; 2001-327840/34.
DR P-PSDB; ABUS3319.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Claim 1; Page 942-943; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence encodes a polypeptide
CC described in the disclosure of the invention
CC
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
Query Match 47.4%; Score 1374.2; DB 5; Length 3443;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy 82 CTACTGCTGCCCGACCACTCCACTCGCGGTCCCGGACGATGAGTGGCTTAC 141
Db 87 CTACTGATTTCCCTGCGCCCTTCTCACTCTGCTCGCATGAGTGGCTTTCG 146
Qy 142 TCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCTTGGCTCGGAGAGCGCCGCGCCG 201
Db 147 CGCGGCGCTCGGCGCCCACTCTCTCGGCGGAGACTA-----GCCGCGCCG 194
Qy 202 TGGGGAAGGCTCATCAAGTGGCTTAACCTGTGATGAAGAGAGAGCGGTGAAGCGGC 261
Db 195 TGGGGAAGGCTGATCAAGTGGCGCAAGCCCTTAATGAAGAGAGCGGTGAAGCGGC 254
Qy 262 ACCATCAAAACACAACTGTGGGACCGGCTACAGTTCTTGAAGACGTTGGGCAAGGCA 321
Db 255 ACCACCAACAGCAACCTGGGACCGGCTACAGTTCTTGAAGACGTTGGGCAAGGCA 314
Qy 322 CCTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGCGCTGTGGCCATCAAGTCA 381
Db 315 CCTACGGGAAGGTGAAGAGGCGGAGAGCTCGGGCGCTGTGGCCATCAAGTCA 374
Qy 382 TCAGAAAGACAAATCAAGATGAGAGCATCTGCTGACATACGAGGAGATTTGA 441
Db 375 TCCGGAAGGACAAATCAAGATGAGAGCATCTGATGACATACGAGGAGATTTGA 434

OY	442	TCATGTCCTTCACTCAACCAACCCCAATCATATGCGATCCATCAATGAATGTTTGAATATGACA	501
Db	435	TCATGTCCTTCACTCAACCAACCCCAATCATATGCGATCCATCAATGAATGTTTGAATATGACA	494
OY	502	GCAAGATTGTGATTGTGATGAGTATGACAGCCGAGCGGATGTGTATGATTATCATCATG	561
Db	495	GCAAGATTGTGATTGTGATGAGTATGACAGCCGAGCGGATGTGTATGATTATCATCATG	554
OY	552	AGCGGCCACGGGTGATGAGCGGGACGCCAGCATTTCTTCCGACAGATCGTGTGCCC	621
Db	555	AGCGGCCACGGGTGATGAGCGGGACGCCAGCATTTCTTCCGAGAGATCGTGTGCCC	614
OY	622	TGCACTACGCGCACCAAGAACGGGATTCGTTCAACCGAATTCACAGCTGGAAAAATCTCTC	681
Db	615	TGCACTATTTGCGCATGAAGACAGATTGTTCACCGAGATCTCAAGTGGAGAACATCTCTC	674
OY	682	TAGATGCCAATGGAACATCAAGATTGCTGACCTTGGGCTCTCCAACTGTATCCATCAAG	741
Db	675	TGAGATGCCAATGGAATATCAAGATTGCTGACCTTGGGCTCTCCAACTGTATCCATCAAG	734
OY	742	GCAAGTTCCTCCAGCGTTCTGTGGAGGCCCTCTAGCGCTTCGCTGAGATATGCAACG	801
Db	735	GCAAGTTCCTCCAGCATTTCTGTGGAGGCCCTCTATGAGCTTCGCAAGATGTCAATG	794
OY	802	GGAAGCCCTATGTGGGCCCAAGGTGTGACAGTGGTCTCTGGGGCTTCTCTGATCATCC	861
Db	795	GGAAGCCCTTATCAAGGCCCAAGGTGTGACAGTGTGCTCTGGGGTCTCTCTGATCATCC	854
OY	862	TGATGTCATGCGACATATGCGCTTGTGACGGGACAGATCATTAACACTGGTGAAGCAATCA	921
Db	855	TGATGTCATGCGACATATGCGCTTGTGATGGGCATGACATTAAGTCTTATGTAACAGATCA	914
OY	922	GTAACGGGGCTTACCGGTAGCCGCCAAGCCGTCCGATGCTGTGGCTGATCCGGTGGC	981
Db	915	GCAACGGGGCTTACCGGGAGCCACTTAACCTCTGATGCTGTGGCTGATCCGGTGGC	974
OY	982	TGTTTATGTGTAACCCCAACCCGCTGGGGCCACCTGGAAGATGTAGCCATGATTTGGTGGG	1041
Db	975	TGTTTATGTGTAACCCCAACCCGCTGGGGCCACCTGGAAGATGTAGCCATGATTTGGTGGG	1034
OY	1042	TCAACTGGGGTTTACCAACCGGAGTCGGGGAAACAAGAAAGCCCTGGTGAAGGTGGGACAC	1101
Db	1035	TCAACTGGGGCTTACCGCAACCGAGTGGGAAGACAAGAGCTTCGATAGAGGTGGGACAC	1094
OY	1102	CTAGTGTGATCTTTGGCCGGGCTTCATGCGCGGACTGTGTACGTGCTCTCGCGCCCC	1161
Db	1095	CTGGGAGTGACTCTGCGCGCGCTCTCAATGCTGATCTGGCTCTCGCGCCCC	1154
OY	1162	TCCTGGAGATGGAACCAAGGTGTGACGTTCTTCAAGCAGACGTCGCGAGGTGGAA	1221
Db	1155	TCCTGGAGAAATGGGGCCAAAGGTGTGACGTTCTTCAAGCAGATGCACTGGTGGGGA	1214
OY	1222	GCACTGTATCTGGGCTGGAAGCGGCAACATTCTCTTAAGAATCTCCGAAAGGAATGACA	1281
Db	1215	GCACCACTCCGCGCTTGAAGGCGCAAGATTGCTCTCAAGAAAGTCCCGCAAGAGATGACA	1274
OY	1282	TGGCTCAAAATCTGCAAGGTGACCGCGGCTGAAGATACGTTCTTCGCGCTGGCAAGACA	1341
Db	1275	TGGCCCAATCTCTCAACATGACACGGCTGATGACATGCGCATGCGCTTGGCAAGACA	1334
OY	1342	GCCTTAAGCTTCCGAAGGCAATTCTCAAGAAAAAGTCTCTTACCTCTGACAGGGAGTAC	1401
Db	1335	ACCTCAAGCTGCAAAAGGGCAATTCTCAAGAAAGATGTACGCTCTGCAAGAGGGTAC	1394
OY	1402	AGAGGAACCTTCAGGAATTCAGACCAGTGTGCTGATATCTCAAGGAGCGCTGTGCTGTG	1461
Db	1395	AGAGGAACCTTCAGGAGCTCAGGCCAATCTCTGCAAGCCAGAGGAGGCTGACC-----	1448
OY	1462	TATCCCTGCTCCCAAGGAAGGCAATCTTAAGAGTCTCGACGCTGAAATCTGTATCT	1521
Db	1449	---CGTGTCTCCCAAGAGGCAATTTCAAGAAACCCCGACAGCGCATGTCTGTGATCT	1505

Qy	1522	ACTCTCTCCAGAGCCCAAGCAGATCTGGGGAACTCTTAAAGCCAGATGATGTGTTTGTA	1581
Db	1506	ACTCTCTCCCGAAGCCAGTAAATCTGGGGAGCTTTGAAAGCAGGGAGACTGTTTGTA	1565
Qy	1582	GTGGGGAGCCCGTGGAGAGAAAGTCTCCACAGGCTTACAGGGCTCTCTCCACCGCAGG	1641
Db	1566	GTGGGGATCCCAAGAGCAGAAAGCTCCGACAGCTTACAGGGCTCTCTCATGCGAAG	1625
Qy	1642	GCATTCTCAAACTCAATGGCAAGTTCTTCCGCAAGCCTTAAAGAGCACTACCCCTAGCA	1701
Db	1626	GCATCTCTCAAACTCAATGGCAAGTTCTTCCGACAGCCTTGAAGCTGCGGGCCCCACA	1685
Qy	1702	CCTTTGGTCCCTTGGACCAATGGCCTCTCTCCACTCTGACAGCCCGGCCACGCTCCCT	1761
Db	1686	CCCTTCGGCTCCCTGGATGAATCTGCCCCACCTCCGCCCTGGGCCCGGACAGCTGACCT	1745
Qy	1762	CAGGGGCTGTAGTGAAGGACAGCATCCGTGCTCCGAGTCCCTTGAACAAATTGACCTTGC	1821
Db	1746	CAGGGGCTGTAGTGAAGGACAGCATCCGTGCTTGAAGTCTTGAACCAAGCTGAGCTTGC	1805
Qy	1822	CTGAACGTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGACAACTGAGCGGGC	1881
Db	1806	CTGAACGGCTCCAGAGAGCCCCCACTGGGGGGCTGTGTCTGTGACAACTGACGGGGC	1865
Qy	1882	TTGAGCAGCCTCTCCCTCAGAG-----GTGGAAGCATGTGTGCGAGATCTT	1929
Db	1866	TTGAGAGAGCCCCCTCAGAGGGCCCTGGAAGCTGCTGAGGGCGCTGCGCAGGATCTT	1925
Qy	1930	TGGGGGATAGTGTGCTTTTCTGTGA CAGACTGCGCAAGAGGTGACTGCGACCTTACAGCAAG	1989
Db	1926	TGGGGGACAGCTGTCTTTTCTGTGA CAGACTGCGCAGAGAGTGA CAGCACTTACCGACAGG	1985
Qy	1990	CCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAAGGAGATGTGCTCCCTAGTATAGGGTAG	2049
Db	1986	CACAGAGGCTGTGCTCAAAAGCTCA CTAAGTGAAGTGAAGGCAATTGCCCCAG--CCCGGTGAG	2044
Qy	2050	GCTCTGAGAGGGTTTGCAAGAGAACCTGGGTGGATTTCTCC	2092
Db	2045	GCTCTCAGATGCAAGCTGTGTGCA CCGAGGGGAGATGCTTTC	2087
RESULT 11			
ADF76964			
ID	ADF76964	standard; cDNA; 3443 BP.	
XX	AC	ADF76964;	
XX	DT	26-FEB-2004 (first entry)	
XX	DE	Novel human secreted and transmembrane protein cDNA Segid 639.	
XX	KM	human; PRO; membrane bound protein; membrane bound receptor;	
XX	KM	cell proliferation; cell migration; cell differentiation;	
XX	KM	mitogenic factor; survival factor; cytotoxic factor;	
XX	KM	differentiation factor; neuropeptide; hormone; cell receptor;	
XX	KM	receptor-ligand interaction; cytoskeletal; chondrocyte; tumour; ss; gene.	
OS		Homo sapiens.	
XX	XX	WO2003072035-A2.	
XX	FN	04-SEP-2003.	
XX	PD	21-FEB-2003; 2003WO-US005241.	
XX	PE	22-FEB-2002; 2002US-0359461P.	
XX	FR		
XX	PA	(GETH) GENENTECH INC.	
XX	PI	Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;	
XX	PI	Williams FM, Wood WI, Wu TD;	
XX	DR	WPI; 2003-721702/68.	

DR P-PSDB; ADF76965.
XX New PRO polypeptides, useful for diagnosing and treating an immune
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
XX diabetes mellitus.

PS Claim 2; SEQ ID NO 639; 918bp; English.

XX This invention relates to novel nucleic acids encoding human PRO secreted
CC and transmembrane proteins. Extracellular proteins play important roles
CC in the formation, differentiation and maintenance of multicellular
CC organisms. The fate of many individual cells (for example proliferation,
CC migration or differentiation) is typically governed by information
CC received from other cells and the immediate environment. The information
CC is often transmitted by secreted polypeptides (for example mitogenic
CC factors, survival factors, cytotoxic factors, differentiation factors,
CC neurotrophins and hormones) which are received and interpreted by diverse
CC cell receptors or membrane bound proteins. These membrane bound proteins
CC and receptors may be of use as pharmaceutical and diagnostic agents, such
CC as in the blocking of receptor-ligand interactions. The current invention
CC provides the amino acid sequences of novel human membrane bound receptors
CC and proteins, along with the cDNA sequences encoding them. The novel
CC proteins of the invention may have cytostatic activities through the
CC stimulation of chondrocytes. The nucleic acids of the invention may be
CC useful for the manufacture of a medicament for diagnosing or treating a
CC tumour in a mammal. In addition, they may be useful for measuring or
CC detecting the expression of a tumour associated gene. The present
CC sequence is a cDNA sequence which encodes a human PRO protein of the
XX invention.

XX Sequence 3443 BF; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 47.4%; Score 1374.2; DB 10; Length 3443;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;

Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGTCGCCGACCTTCACCTGCGGTCGCCCAACCTGAGTCGGGCTTAC 141
DB 87 CTACTGTCGCCGACCTTCACCTGCGGTCGCCCAACCTGAGTCGGGCTTAC 146
QY 142 TCCAGCGCCGAGCGAGGCTTCCTGCGGCTTCGCGCTCGAGAGCGCCGCGC 201
DB 147 CGCGGCGCTCGCGGCGCCACTCTCTCGCGCGGAGAGCTA-----GCCGCGCGC 194
QY 202 TGGCGGAGCGGCTCATCAATCGCTTAACTCTGATGAGAGACAGCGGTGAAGGCGC 261
DB 195 TGGCGGAGCGGCTCATCAATCGCTTAACTCTGATGAGAGACAGCGGTGAAGGCGC 254
QY 262 ACCATCAAAACAAACCTGCGGAGCGGCTGAGTTCTTGAGAGCGCTGGGCAAGGCA 321
DB 255 ACCATCAAAACAAACCTGCGGAGCGGCTGAGTTCTTGAGAGCGCTGGGCAAGGCA 314
QY 322 CTACCGGAGAGGTGAAGAGCAAGAGAGCTCGGGCGCTCTGTGGCCATCAAGTCA 381
DB 315 CTACCGGAGAGGTGAAGAGCAAGAGAGCTCGGGCGCTCTGTGGCCATCAAGTCA 374
QY 382 TCGAGAAAGCAAAATCAAAAGATGAGAGAGTCTGTCACATACGAGGAGATTGAGA 441
DB 375 TCGAGAAAGCAAAATCAAAAGATGAGAGAGTCTGTCACATACGAGGAGATTGAGA 434
QY 442 TCATGTCCTCACTCAACACCCCAACATTCATTCATTCATGAGAGTTGAGAAATGCA 501
DB 435 TCATGTCCTCACTCAACACCCCAACATTCATTCATTCATGAGAGTTGAGAAATGCA 494
QY 502 GCAAGATGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 495 GCAAGATGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 554
QY 562 AGCGGCAACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
DB 555 AGCGGCAACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 614

QY 622 TGCACTATGCGACCAAGAGCGGATCGTTACCGGAGATCTCAACCTGAGAAACTCTTC 681
DB 615 TGCACTATGCGACCAAGAGCGGATCGTTACCGGAGATCTCAACCTGAGAAACTCTTC 674
QY 682 TAGATGCGAATGAGAAATCAAGATGAGATGAGATGAGATGAGATGAGATGAGATG 741
DB 675 TAGATGCGAATGAGAAATCAAGATGAGATGAGATGAGATGAGATGAGATGAGATG 734
QY 742 GCAAGTCTCTCAAGACCTGTCGAGAGCCCTCTCAAGCTCTGAGATGATGATGATG 801
DB 735 GCAAGTCTCTCAAGACCTGTCGAGAGCCCTCTCAAGCTCTGAGATGATGATGATG 794
QY 802 GGAAGCCCTATGTCGAGGCGGAGAGTGAACAGTGTCTCTGAGGCTTCTCTGATC 861
DB 795 GGAAGCCCTATGTCGAGGCGGAGAGTGAACAGTGTCTCTGAGGCTTCTCTGATC 854
QY 862 TGGTGCATGAGCAACATGCTTGAAGGAGAGATGATGATGATGATGATGATGATG 921
DB 855 TGGTGCATGAGCAACATGCTTGAAGGAGAGATGATGATGATGATGATGATGATG 914
QY 922 GTAAAGGAGCTTACCGTGAAGCGGCGCAAGCTGTCGATGCTGTCGATGCTGTCG 981
DB 915 GCAAGGAGGCTTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 974
QY 982 TGTATATGATGAAACCCCAAGCGGCGCAACATGAGATGATGATGATGATGATG 1041
DB 975 TGTATATGATGAAACCCCAAGCGGCGCAACATGAGATGATGATGATGATGATG 1034
QY 1042 TCAACTGAGGCTTACCAACCGGAGTGGGAGACAGAGAGCCCTGTCGAGGCTGAG 1101
DB 1035 TCAACTGAGGCTTACCAACCGGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1094
QY 1102 CTAGTGTGATCTTGGCGGCGGCTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1161
DB 1095 CTAGTGTGATCTTGGCGGCGGCTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1154
QY 1162 TCTGAGAGATGAGAGCGGAGTGTGAGCTTCTTCAAGACAGAGGTCGAGGAGTGA 1221
DB 1155 TCTGAGAGATGAGAGCGGAGTGTGAGCTTCTTCAAGACAGAGGTCGAGGAGTGA 1214
QY 1222 GCACTGATCTGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1281
DB 1215 GCACTGATCTGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1274
QY 1282 TGGCTCAAAATGTCAGAGTGAACCGGCTGAGGATGATGATGATGATGATGATG 1341
DB 1275 TGGCTCAAAATGTCAGAGTGAACCGGCTGAGGATGATGATGATGATGATGATG 1334
QY 1342 GCTTAAAGCTTCCGAAAGGAGATCTTCAAGAAAGGCTCTTCAAGCTGTCAGAGGAGTAC 1401
DB 1335 ACCTCAAGCTGTCAGAAAGGAGATCTTCAAGAAAGGAGTCTTCAAGCTGTCAGAGGAGTAC 1394
QY 1402 AGAGAGACCTTCAAGAACTGAGCGGCTGATCTTCAAGAGGAGCTGTCCTGCTG 1461
DB 1395 AGAGAGACCTTCAAGAACTGAGCGGCTGATCTTCAAGAGGAGCTGTCCTGCTG 1448
QY 1462 TATCCCTGCTCCCAAGAAAGGAGTCTTCAAGAAAGTCTTCAAGAGGAGTCTG 1521
DB 1449 ---CGTCTCTCCCAAGAAAGGAGTCTTCAAGAAAGTCTTCAAGAGGAGTCTG 1505
QY 1522 ACTCTCTTCAAGAGCGGAGGAGTCTGAGGAACTTCAAGAGCGGAGTCTGAGGAGTCTG 1581
DB 1506 ACTCTCTTCAAGAGCGGAGGAGTCTGAGGAACTTCAAGAGCGGAGTCTGAGGAGTCTG 1565
QY 1582 GTGGGAGCCCGTGGAGCAAGAGTCTTCAAGAGGCTTCAAGAGGCTTCTTCAAGCGCAAG 1641
DB 1566 GTGGGAGTCCCAAGAGCAAGAGTCTTCAAGAGGCTTCAAGAGGCTTCTTCAAGCGCAAG 1625
QY 1642 GCATTCTCAAACTCAATGAGGAGGAGTCTTCAAGAGGCTTCAAGAGGAGGAGGAGGAG 1701
DB 1626 GCATTCTCAAACTCAATGAGGAGGAGTCTTCAAGAGGCTTCAAGAGGAGGAGGAGGAG 1685
QY 1702 CTTTGGCTCTTGGAGCAAACTGAGCTTCTTCCATCTGAGAGCGGCGGCGCTT 1761

Db 1686 CTTTGGCTCCCTGGATGAATCGCCACCTCCGCCCTGGCCCGGCGCAGCCGACCT 1745
Qy 1762 CAGGGGCTGTGATGAGAGACGACATCTGTCTCCGAGTCTTTGACCAATTTGACTTG 1821
Db 1746 CAGGGGCTGTGAGGAGAGACGACATCTGTCTCTGAGTCTTTGACCACTGACTTGAC 1805
Qy 1822 CTGAACGCTTTCCCGAAGACCCCACTGAGGGGCTGTGTCTGTGTGACAACTTGAGGGGC 1881
Db 1806 CTGAACGCTTTCCCAAGAGCCCACTGAGGGGCTGTGTGTGTGTGACAACTTGAGGGGC 1865
Qy 1882 TTGAGCAGCTCTCCCTCAGAG------GTCTGAAGCGATGTGTGAGAGATCTT 1929
Db 1866 TTGAGAGAGCCCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1925
Qy 1930 TGGGGGATGAGTGTCTTTCTCTGACAGACTGCAAGAGTGTGAGAGGAGGAGGAGGAG 1989
Db 1926 TGGGGGAGAGGCTTTTCTCTGACAGACTGCAAGAGTGTGAGAGGAGGAGGAGGAGGAG 1985
Qy 1990 CCCTAGAGATCTGTCTCAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2049
Db 1986 CACTGAGAGGCTGTCTCAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2044
Qy 2050 GCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2092
Db 2045 GCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2087

RESULT 12
ADP81952
ID ADP81952 standard; DNA; 3443 BP.
AC ADF81952;
XX
DT 26-FEB-2004 (first entry)
XX
DE Leukemia-related DNA sequence #2508.
XX
KW Cytostatic; Gene therapy; Leukemia; ss.
OS Unidentified.
XX
PN WO2003039443-A2.
PD 15-MAY-2003.
XX
PF 04-NOV-2002; 2002WO-EP012303.
XX
PR 05-NOV-2001; 2001EP-00126244.
PR 30-APR-2002; 2002EP-00009758.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
PA (HAFE/) HAFERLACH T.
PA (SCHO/) SCHOCH C.
PA (KERN/) KERN W.
XX
PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schultze S, Dugas M,
PI Bils R, Broers B, Mergenthaler S;
XX
DR WPI; 2003-505037/47.
XX
PT Determining the subtype of leukemia cells and whether a patient sample
PT contains leukemia cells or other cells, useful for treating leukemia,
PT comprises determining the expression profile of a group of markers in a
PT patient sample.
XX
PS Disclosure; SEQ ID NO 2508; 2938bp; English.
XX
CC The present invention relates to a method (M1) for determining the
CC subtype of leukemia cells and whether a patient sample contains
CC leukemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for

CC determining the presence of leukemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukemia.
XX
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
Query Match 47.4%; Score 1374.2; DB 10; Length 3443;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy 82 CTACTGTGCGCCGACCCATCTCCAGTCCGAGTCCCGCCGACCACTGAGTCCGCTTAC 141
Db 87 CTACTGATTCCTCGCGCCGCTTGTCTACCTCGTGTGCGCAATGAGATCTCGTTTTCG 146
Qy 142 TCCAGCGCCCGAGCGAGGCTTCTCGGCTTCGCGCTTCGCGCTTCGAGAGCGCCGCGC 201
Db 147 CGGCGCCCTCGGCGCCCATCTCGCGCCGCGCAAGTTA-----GCCCGCGCGC 194
Qy 202 TGGGCGAGCGGCTCATCAATCGCTTAAACCTGTGATGAGAGAGAGGAGGAGGAGGAGGAG 261
Db 195 TGGGCGAGGAGGCTGATCAATCGCTTAAACCTGTGATGAGAGAGGAGGAGGAGGAGGAGGAG 254
Qy 262 ACCATCAAAAACAACCTGCGGACCGCTTACAGATTTCTGAGAGCGCTGAGAGGAGCA 321
Db 255 ACCACCAACAAGCAACCTGCGGACCGCTTACAGATTTCTGAGAGCGCTTACAGAGGAGCA 314
Qy 322 CTTACGGGAGGAGTGAAGAGGAGCAAGAGAGCTTGGGGGCTGTGTGGCCATCAATGCA 381
Db 315 CTTACGGGAGGAGTGAAGAGGAGCGGAGAGGAGCTTGGGGGCTGTGTGGCCATCAATGCA 374
Qy 382 TCAAGAAAGACAAATCAAGATGAGAGGAGTCTGCTGACATACGAGGAGGAGGAGTGTGAG 441
Db 375 TCCGAGAGAGCAAAATCAAGATGAGAGGAGTCTGCTGACATACGAGGAGGAGGAGTGTGAG 434
Qy 442 TCATGCTTCTCAACCAACCCCACTATTCGATCCATGAGAGTGTGAGAGTGTGAGAGTGTGAG 501
Db 435 TCATGCTATCATCAACCAACCCCACTATTCGATCCATGAGAGTGTGAGAGTGTGAGAGTGTGAG 494
Qy 502 GCAAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 495 GCAAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
Qy 562 AGGCGGACAGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
Db 555 AGGCGGACAGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 614
Qy 622 TGCACTACTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
Db 615 TGCACTATTCGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 674
Qy 682 TAGATGCCAATGGAACATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
Db 675 TGATGCCCAATGGAATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
Qy 742 GCAAGTTCCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
Db 735 GCAAGTTCCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 794
Qy 802 GGAAGCCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 861
Db 795 GGAAGCCATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 854
Qy 862 TGGTGCATGGCACCATGCTCTTTGACGGGAGGAGTATATAACATGTTGAGAGCAATATCA 921
Db 855 TGGTGCATGGCACCATGCTCTTTGAGGGGAGTATATAAGATCTAGTGAACAATATCA 914
Qy 922 GTAAAGGAGGCTTACGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
Db 915 GCAAGGAGGCTTACGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 974
Qy 982 TGTAAATGAGAACCCACCCGTCGGGAGGAGCACTGAGAGATGATGAGCAATGATGAGGAG 1041
Db 975 TGTAAATGAGAACCCACCCGTCGGGAGGAGCACTGAGAGATGATGAGCAATGATGAGGAG 1034

QY 1042 TCAACTGGGTTACCAACCGGAGTCGGGGAACAGAAAGCCCTGCGTGAAGGCTGGGACC 1101
 DB 1035 TCAACTGGGTTACCAACCGGAGTCGGGGAACAGAAAGCCCTGCGTGAAGGCTGGGACC 1094
 QY 1102 CTGAGTGAATCTTGGGCGGGGCTCCATGGGCGAGCTGGTGAAGTGGCTCCCTGGGCCCC 1161
 DB 1095 CTGAGTGAATCTTGGGCGGGGCTCCATGGGCGAGCTGGTGAAGTGGCTCCCTGGGCCCC 1154
 QY 1162 TCCTGGAATGAGCAAGGTTGCAAGCTTCTTCAAGCAGACAGTGCAGGAGGATGGA 1221
 DB 1155 TCCTGGAATGAGCAAGGTTGCAAGCTTCTTCAAGCAGACAGTGCAGGAGGATGGA 1214
 QY 1222 GCACTGTAAGTGGGCTGAGCGGCAACATCTTCAAGAGTCCCGAAGGAGATGCA 1281
 DB 1215 GCACTGTAAGTGGGCTGAGCGGCAACATCTTCAAGAGTCCCGAAGGAGATGCA 1274
 QY 1282 TGGCTCAAAATCTGCAAGTGAACCCGGTGAAGTACCTCTTCCGCTGGCAAGACA 1341
 DB 1275 TGGCTCAAAATCTGCAAGTGAACCCGGTGAAGTACCTCTTCCGCTGGCAAGACA 1334
 QY 1342 GCCTTAAGCTTCCGAAAGGCAATCTTCAAGAAAGTCTTACCTCGTCAAGGAGGTAC 1401
 DB 1335 ACCTCAAGCTTCCGAAAGGCAATCTTCAAGAAAGTCTTACCTCGTCAAGGAGGTAC 1394
 QY 1402 AGAGAGACCTTCAAGAACTCAAGACCGGTGCTGATTAAGGAGGAGCTGCTGCTG 1461
 DB 1395 AGAGAGACCTTCAAGAACTCAAGACCGGTGCTGATTAAGGAGGAGCTGCTGCTGCTG 1448
 QY 1462 TATCCCTGCTCCCAAGAAAGGCAATCTTCAAGAAAGTCTTCAAGGAGGAGCTGCTGCTG 1521
 DB 1449 TATCCCTGCTCCCAAGAAAGGCAATCTTCAAGAAAGTCTTCAAGGAGGAGCTGCTGCTG 1505
 QY 1522 ACTCCTCTCCAGAGCCCGAGAGTCTGGGGAATCTTGAAGCCAGTGAATGTTTGA 1581
 DB 1506 ACTCCTCTCCAGAGCCCGAGAGTCTGGGGAATCTTGAAGCCAGTGAATGTTTGA 1565
 QY 1582 GTGGGAGACCCGCTGAGAGAGAGTCTTCAAGGCTTCAAGGCTCTCTCCACCGCAGG 1641
 DB 1566 GTGGGAGATCCCAAGAGAGAGAGTCTTCAAGGCTTCAAGGCTCTCTCCACCGCAGG 1625
 QY 1642 GCATTCTCAAACTCAATGGCAAGTCTTCCGCAAGCCTTGAAGGCACTAACCCCTAGCA 1701
 DB 1626 GCATTCTCAAACTCAATGGCAAGTCTTCCGCAAGCCTTGAAGGCACTAACCCCTAGCA 1685
 QY 1702 CCTTGGCTCCCTGAGCAACTGAGCTCTCCACCTCCAGCCGCGCCAGCCGCT 1761
 DB 1686 CCTTGGCTCCCTGAGCAACTGAGCTCTCCACCTCCAGCCGCGCCAGCCGCT 1745
 QY 1762 CAGGAGCTGTAAGTGAAGCAGATCTGCTCCGAGTCTTGAACCAATTGAGCTTGC 1821
 DB 1746 CAGGAGCTGTAAGTGAAGCAGATCTGCTCCGAGTCTTGAACCAATTGAGCTTGC 1805
 QY 1822 CTGAAGCTTCTCCGAAACCCCACTGAGGAGGCTGTGTCTGTGACAACTGAGGAGG 1881
 DB 1806 CTGAAGCTTCTCCGAAACCCCACTGAGGAGGCTGTGTCTGTGACAACTGAGGAGG 1865
 QY 1882 TTGAGGAGCTTCCCTGAGAG-----GTGGAACCAATGTGTGAGAGAAATCT 1929
 DB 1866 TTGAGGAGCTTCCCTGAGAG-----GTGGAACCAATGTGTGAGAGAAATCT 1925
 QY 1930 TGGGAGATGAGTGTCTTCTTCTGACAGACTCCAAAGAGTGAAGTGAAGCTTACAGACAG 1989
 DB 1926 TGGGAGATGAGTGTCTTCTTCTGACAGACTCCAAAGAGTGAAGTGAAGCTTACAGACAG 1985
 QY 1990 CCCTGAAGATGTCTCAAGCTCAAGCTGAGAGAGGAGAGATGTGCTTATGTGGGTAG 2049
 DB 1986 CACTGAGAGGTGTCTCAAGCTCAAGCTGAGAGAGGAGAGATGTGCTTATGTGGGTAG 2044
 QY 2050 GCTCTGAGAGGTTTGAAGAGAACCTGGGTCGAGATTCCTCC 2092
 DB 2045 GCTCTGAGATGAGTGTGAGAACCTGGGTCGAGATTCCTCC 2087

RESULT 13
 ADO20171
 ID ADO20171 standard; cDNA; 3443 BP.
 XX
 AC ADO20171;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human PRO polynucleotide #540.
 XX
 KW Human; PRO; gene; ss; immune related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KW renal disease; demyelinating disease; central nervous system;
 KW peripheral nervous system; demyelinating polyneuropathy;
 KW Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polyneuropathy.
 XX
 OS Homo sapiens.
 XX
 PN WO2004043361-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 06-NOV-2003; 2003WO-US035268.
 XX
 PR 08-NOV-2002; 2002US-0425235P.
 XX
 PA (GENTH) GENTECH INC.
 XX
 PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WJ, Wu TD;
 XX
 DR WPI; 2004-420067/39.
 DR P-PSDB; ADO20172.
 XX
 PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO8638 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 XX
 PS Claim 1; SEQ ID NO 1136; 1731bp; English.
 XX
 CC The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.
 XX
 SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;
 Query Match 47.4%; Score 1374.2; DB 12; Length 3443;
 Best Local Similarity 81.6%; Pred. No. 1.1e-293;
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
 QY 82 CTACTGTCGCCGACCACTTCACTTGCAGTCTCCCGACCATGAGTGGCTTAC 141
 DB 87 CTACTGATTCCTCCGCGCCCTTCTCACTCTGCTGCGCATGAGTGTGTTTCG 146
 QY 142 TCCAGGCGCCGAGCAGAGTCTCCGCGCTCCGCGCTGCGAGAGCGCCGCGCGC 201
 DB 147 CGCGGCGCTCCGCGCCCACTCTCCGCGGAGAGCTA-----GCCGCGCGC 194
 QY 202 TGGCGAGCGGCTCATCAAGTGGCTTAACCTTGAATGAAGACAGGCGGTGAAGCGGC 261

Db 195 TGGGGAGAGGCTATCAAGTCGCGCAAGCCCTTATAGAGAGCGGTTGAGCGCC 254
Qy 262 ACCATCACAACACACACTGCGGACCGCTACGAGTTCTTGAGAGCCGTGGGCAAGGCA 321
Db 255 ACCACCAAGACCAACCTGCGGACCGCTACGAGTTCTTGAGAGCCGTGGGCAAGGCA 314
Qy 322 CCTACGGGAGGTTGAGAGGACGAGAGAGCTCGGGGCTGTGTGGCCATCAAGTCA 381
Db 315 CCTACGGGAGGTTGAGAGGACGCGGGAGAGCTCGGGGCTGTGTGGCCATCAAGTCA 374
Qy 382 TCAGAGAGACAAATCAAGATGAGAGAGTGTCTGCATACATACGAGGAGATTGAGA 441
Db 375 TCCGAGAGAGCAAAATCAAGATGAGAGAGTGTCTGCATACATACGAGGAGATTGAGA 434
Qy 442 TCATGTCTTCACTCAACCAACCCCAATCATTTGCGATCCATGAAGTTTGAAGATAGA 501
Db 435 TCATGTATCATCACTCAACCAACCCCAATCATTTGCGATCCATGAAGTTTGAAGACAGA 494
Qy 502 GCAAGATTGTGATGTCTATGAGATATGCGACCGAGCGGAGTGTATATATCAATCAAGT 561
Db 495 GCAAGATTGTGATGTCTATGAGATATGCGACCGAGCGGAGCTTTATGATCAATCAAGCG 554
Qy 562 AGCGGACACGAGCTGATGAGGCGGACGCGACAGATTTCTCCGACAGATCGTGTGCGCC 621
Db 555 AGCGGACACGAGCTGATGAGGCGGACAGCTTATGCGATTTCTCCGACAGATCGTGTGCG 614
Qy 622 TGCACTACGCGACCAAGACGGAGTCTTCAACCGAGATCTCAAGCTGAGAAACATCTTC 681
Db 615 TGCACTATGCGCATCAAGACAGAGTGTGTCACCGAGATCTCAAGCTGAGAAACATCTCT 674
Qy 682 TAGATGCGCAATGAGAAACATCAAGATGTGATCTTGGCTCTCCACCTGTACCAAG 741
Db 675 TAGATGCGCAATGAGAAACATCAAGATGTGATCTTGGCTCTCCACCTGTACCAAG 734
Qy 742 GCAAGTCTCTCAAGCTTCTGTGGAGCCCTCTCTACGCTCGCTGAGATGATGACAG 801
Db 735 GCAAGTCTCTGCGACATTTCTGTGGAGCCCTCTCTATGCTCTCGCGAGATTTGTCAATG 794
Qy 802 GGAAGCCCTTATGTGGGCGCCAGAGGTGACAGCTGTCTCTGGGCGTTCTCTGTACATCC 861
Db 795 GGAAGCCCTTACACAGGCGCCAGAGGTGACAGCTGTCTCTGGGCGTTCTCTGTACATCC 854
Qy 862 TGTGTGATGAGCAACATGCTCTTGAACGGGACAGATCATTAACATCTGTGAGCAATCA 921
Db 855 TGTGTGATGAGCAACATGCTCTTGAATGGGACAGACATTAAGATCTGTGAGCAATCA 914
Qy 922 GTAACGGGCGCTTACCGTGAAGCGCCCAAGCGGTCCGATGCTGTGGCCGTATCCGGTGGC 981
Db 915 GCAACGGGCGCTTACCGGAGGACCTTAACCTCTGTATGCTGTGTGGCTGTATCCGGTGGC 974
Qy 982 TGTATATGTGAACCCCAACCGGTGGGCGACATGAGAGATGAGCCAGTCAATGTGGG 1041
Db 975 TGTATATGTGAACCCCAACCGGCGGCGACCTGTGAGAGATGTGGCGCATCTGTGTGG 1034
Qy 1042 TCAACTGGGGTTTACACCAACCGAGTCCGGAGAACGAAACCTTCGCTGAGGGGTGGCAAC 1101
Db 1035 TCAACTGGGGCTTACGCAACCGAGTGGGAGAGAGGCTCCCAATGAGGGTGGCAAC 1094
Qy 1102 CTAGTGTGATCTTGTGGCGGGCGCTCAATGGGAGCTGGTTACGTGTGCTCTCGGCGCC 1161
Db 1095 CTAGTGTGATCTTGTGGCGGGCGCTCAATGGGAGCTGTGCTGTGGCGCTCTCGGCGCC 1154
Qy 1162 TCCTGAGAGATGAGCAAGGTGTGCACTTCTTCAAGCAGCAGTCCGCGGAGGTGAGAA 1221
Db 1155 TCCTGAGAGATGAGCAAGGTGTGCAAGTCTTCTTCAAGCAGCAGCAGTGTGTGGGAGAA 1214
Qy 1222 GCACTGTACTGCTGGCTGAGAGCGCAACATTTCTTTAAGAACTCCGAAAGGAGATGACA 1281
Db 1215 GCACTGTACTGCTGGCTGAGAGCGCAACATTTCTTTAAGAACTCCGCAAGGAGATGACA 1274
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATTAAGTCTTCTGTGCGCTGTGGCAAGCA 1341

Db 1275 TGGCCAGTCTCTCCAGATGACACGGCTATGACATCGCCCATCGCCCTGGCAAGACA 1334
Qy 1342 GCCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCTTCACTGTCTGAGGGAGTAC 1401
Db 1335 ACCTCAAGCTGCCAAAGGGCAATTTCTCAAGAAAAGTGTGAGCTGTCTGCAAGAGGGTAC 1394
Qy 1402 AGAGGACCCCTGAGGAACTGAGACCGGTGTGATATCTCAGGGGACCTGTCTCTGTG 1461
Db 1395 AGAGGACCCCTGAGGAACTGAGACCGGTGTGATATCTCAGGGGACCTGTCTCTGTG 1448
Qy 1462 TATCCCTGTCTCCAGGAAAGGCAATCTTAAAGATCTGACACAGCTGAAATCTGTACT 1521
Db 1449 ---GCTGTCTCCCAAGAGGCAATTTCTCAAGAAACCCCAAGCGGAGCTGTGACT 1505
Qy 1522 ACTCTCTCCAGAGCCGAGGAGTCTGGGAACTTTAGACGCCAGTGAATGTGTTGTA 1581
Db 1506 ACTCTCTCCGAGCCGAGGAAATCTGGGAGGCTTTGAGACGAGGCGAGCGTGTGTA 1565
Qy 1582 GTGGGACCCCTGAGGACAGAAATCTGACAGGCTTCAAGGCTCTCTCTCAACCGCAAG 1641
Db 1566 GTGGGAGTCCCAAGAGACAGAAAGCTCCGCAAGCTTCAAGGCTGTCTCTCATGCAAG 1625
Qy 1642 GCATTTCAAACTCAATGGCAAGTTCTCCGACAGCTTGAAGAGCACTAACCCCTAGCA 1701
Db 1626 GCATTTCAAACTCAATGGCAAGTTCTCCGACAGCTTGAAGCTGTGCGGCGCCCACTCA 1685
Qy 1702 CTTTGTGCTCTCTGAGCAACTGTGCTCTCTCCATCTGAGACCGCGGCGCAAGCGCCCT 1761
Db 1686 CTTTGTGCTCTCTGAGCAACTGTGCTCTCTCCATCTGAGACCGCGGCGCAAGCGCCCT 1745
Qy 1762 CAGGGGCTGTGAGAGGAGACAGCATCTGTCTCTCCAGTCTCTTGAACCAATTTGACTTGC 1821
Db 1746 CAGGGGCTGTGAGAGGAGACAGCATCTGTCTCTCTGAGTCTCTTGAACCAAGCTTGC 1805
Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGACCAACTGTGGGGGCG 1881
Db 1806 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGACCAACTGTGGGGGCG 1865
Qy 1882 TTGAGCAGCTCTCCCTCAAGG-----GTCTGAAGCAGTGTGTGGCAGGAATCT 1929
Db 1866 TTGAGGAGCCCTCCCTCAAGGCGCTGAGAGCTGTGAGAGCGTGTGGCGGAGATCTT 1925
Qy 1930 TGGGGGATAGCTGTTTCTGTACACATCTGCCAAGAGTGTGATGACGCTTACAGACAG 1989
Db 1926 TGGGGGACAGCTGTTTCTGTGACACATCTGCCAAGAGTGTGATGACGCTTACAGACAG 1985
Qy 1990 CCTGAGGATCTGTCAAAAGCTGACGTGAGGAAAGGAGATGTGGCCCTAATGATGGGTAG 2049
Db 1986 CACTGAGAGGTCTGTCAAAAGCTGACGTGAGGAGTGTGGCATTTGCCAG--CCGCTGAG 2044
Qy 2050 GCTCTGAGAGGTTTTCAGAGGAAACCTGTGGATCGATTCCTCC 2092
Db 2045 GCTCTGAGATGCAAGCTGTGTGACCCCGAGGGGAGATGCTTTC 2087

RESULT 14
ADJ96554
ID ADJ96554 standard; DNA; 3463 BP.
XX
AC ADJ96554;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human calcium/calmodulin-dependent protein kinase Nnak2 DNA Segid 11.
XX
KW gene, ds, kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW calcium/calmodulin-dependent protein kinase; Nnak2.
XX
OS Homo sapiens.

OS 68. Location/Qualifiers
FH Key replace(1670,c)
FT variation /tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation /replace(1727,a)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism"
PN MO2004006838-A2.
XX 22-JAN-2004.
XX 15-JUL-2003; 2003WO-US021730.
XX 15-JUL-2002; 2002US-0395632P.
XX (SUGEN-) SUGEN INC.
XX PI Whyte D, Manning G, Caenepeel S;
XX WPI: 2004-122753/12.
XX P-PSDB: ADJ96620.
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX Example 1: SEQ ID NO 11; 366bp; English.
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTKs and STKs),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic, and inflammatory disorders. Accordingly, they exhibit
CC cytoprotective, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
XX
SQ Sequence 3463 BP; 732 A; 1057 C; 919 G; 755 T; 0 U; 0 Other;
Query Match 47.4%; Score 1374.2; DB 12; Length 3463;
Best Local Similarity 81.6%; Pred. No. 1.1e-293;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
QY 82 CTACTGCTGCCGACCACTCACTCGCGGTCCCGACACATGAGTCGTGGCTTAC 141
DB 148 CTATTGATTCCTCCGCGCCCTTGCTCACTCTGCTGCGCATGAGTCGTGTTTCG 207
QY 142 TCCAGCGCCGAGCGAGGCTCCCTCGGCTCGCCCTGCGAGCGCGCGCGCG 201
DB 208 CGCGGCGCTCGGCGCCCACTCTCGCGCGAGCTA-----GCCGCGCGCG 255
QY 202 TGGCGGAGCGGCTCATCAAGTCGCTTAACTCTGATGAGAGAGCGGCTGAAGCGGC 261
DB 256 TGGCGGAGAGGCTGATCAAGTCGCGCAAGCCCTTAATGAAGAGCGGCTGAAGCGGC 315
QY 262 ACCATCAAAACAACCTCGCGGACCGCTACAGATTCTTGAGACGCTGGGACAGGCA 321
DB 316 ACCACCAAAACAACCTCGCGGACCGCTACAGATTCTTGAGACGCTGGGACAGGCA 375
QY 322 CCTACGGAAGGTGAAGAGGACGAGAGAGCTCGGCGCTCTGTGTGCTCAAGTCA 381
DB 376 CCTACGGAAGGTGAAGAGGCGGAGAGAGCTCGGCGCTCTGTGTGCTCAAGTCA 435
QY 382 TCAGGAAGACAAATCAAGATGACGAGATCTGCTCACTACGAGGAGGATTTGA 441

DB 436 TCCGAAAGACAAATCAAGATGACGAGATCTGATGACATACGAGGAGATTGCA 495
QY 442 TCATGCTTTCATCAACACCCCAATCATTTGCTCATCAGTAAGTGTGTAATAGCA 501
DB 496 TCATGCTTTCATCAACACCCCAATCATTTGCTCATCAGTAAGTGTGTAATAGCA 555
QY 502 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 556 GCAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
QY 562 AGCGGCAACCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 621
DB 616 AGCGGCAACCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 675
QY 622 TGCACTACTGCTCAACAGACGAGATCTGTCACGAGATCTCAAGCTGGAACATCTTC 681
DB 676 TGCACTACTGCTCAACAGACGAGATCTGTCACGAGATCTCAAGCTGGAACATCTTC 735
QY 682 TAGATGCCAATGGAACATCAAGATGCTGATCTTGGCTCTGCAACCTGTACCAAG 741
DB 736 TAGATGCCAATGGAACATCAAGATGCTGATCTTGGCTCTGCAACCTGTACCAAG 795
QY 742 GCAAGTCTCTCAACAGCTTCTGTGAGAGGCTCTCTACGCTTCCGCTGATGATGACAG 801
DB 796 GCAAGTCTCTCAACAGCTTCTGTGAGAGGCTCTCTACGCTTCCGCTGATGATGACAG 855
QY 802 GGAAGCCCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 861
DB 856 GGAAGCCCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 915
QY 862 TGGTGCATGAGCAACATGCTTTGACGAGGAGATCATTAACAATGAGGAGCAATCA 921
DB 916 TGGTGCATGAGCAACATGCTTTGACGAGGAGATCATTAACAATGAGGAGCAATCA 975
QY 922 GTAAAGGAGCTTACCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
DB 976 GTAAAGGAGCTTACCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
QY 982 TGTAAATGATGAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
DB 1036 TGTAAATGATGAACCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095
QY 1042 TCAACTGGGCTTACACACCGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
DB 1096 TCAACTGGGCTTACACACCGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1155
QY 1102 CTAGTGTGATTTGGCGCGGCTCTCATGCGGAGTGTGATGATGATGATGATGATG 1161
DB 1156 CTAGTGTGATTTGGCGCGGCTCTCATGCGGAGTGTGATGATGATGATGATGATG 1215
QY 1162 TCCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
DB 1216 TCCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1275
QY 1222 GCACTGTACTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1281
DB 1276 GCACTGTACTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335
QY 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTTGGCAAGACA 1341
DB 1336 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCTCTTCTGCGCTTGGCAAGACA 1395
QY 1342 GCGTTAAGCTTCCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1401
DB 1396 GCGTTAAGCTTCCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1455
QY 1402 AGAGGAGCCCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1461
DB 1456 AGAGGAGCCCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
QY 1462 TATCTGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1521
DB 1510 ---CCTGCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1566

QY 1522 ACTCTCTCCAGAGCCGAGAGTGTGGGGAACTTTAGAGCCAGTGAATGTTTNGA 1581
DB 1567 ACTCTCTCCGAGCCGAGTAACTTGGGGAGCTTTGAGGAGGACGACGTGTTTGA 1626
QY 1582 GTGGGGACCCCGTGGAGAGAGTCTCCACAGGCTTGAGGGCTCTCTCCACGCAAG 1641
DB 1627 GTGGGGATCCCAAGAGAGAGAGCTCCGCAAGTTCAGGGCTGTCTCCATGCAAG 1686
QY 1642 GCATTCCTCAAACTCAATGGCAAGTTCCTCCGCAAGCTTTGAAAGGACCTACCTTACA 1701
DB 1687 GCATTCCTCAAACTCAATGGCAAGTTCCTCCGCAAGCTTTGAAAGTCTCCGCAAG 1746
QY 1702 CCTTGGCTCCCTGAGCAAACTGGGCTCTCCAGTCTCCGAGTCTTTGACCAATTGACTTGC 1761
DB 1747 CCTTGGCTCCCTGAGCAAACTGGGCTCTCCAGTCTCCGAGTCTTTGACCAAGTCTTGC 1806
QY 1762 CAGGGGCTGTGAGTGAAGAGACAGATCTGCTCCGAGTCTTTGACCAATTGACTTGC 1821
DB 1807 CAGGGGCTGTGAGTGAAGAGAGACAGATCTGCTCCGAGTCTTTGACCAAGTCTTGC 1866
QY 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAAACTTGAAGGGC 1881
DB 1867 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTGTGTGTGAGCAAACTTGAAGGGC 1926
QY 1882 TTGAGCAGCCTCCCTCAGAG-----GTGGAAGGATGTGTGAGGAGAACTCT 1929
DB 1927 TTGAGAGGCCCCCTCAGAGGCCCCCTGAGAGTCTCTGAGGCTGTGGGAGAGAACTCT 1986
QY 1930 TGGGGGATAGCTGCTTTTCTTCTGACAGCTGCAAGAGTGTGACAGCTTACAGCAAG 1989
DB 1987 TGGGGGACAGCTGCTTTTCTTCTGACAGCTGCAAGAGTGTGACAGCTTACAGCAAG 2046
QY 1990 CCTTGAAGATCTGTCTCAAGCTCAGCTGAGAGAGAGAGTGTGCTTGAATGAGGTAG 2049
DB 2047 CACTGAGAGGCTGCTCAAGCTCAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2105
QY 2050 GCTCTGAGAGGTTTGTGAGAGAACCTGTGGTCCGATTTCTCC 2092
DB 2106 GCTCTGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2148

RESULT 15

ID ADM43851 standard; cDNA; 2501 BP.

XX AC ADM43851;

XX DT 03-JUN-2004 (first entry)

XX DE Novel human arginine-rich protein cDNA #215.

XX KW ss; gene; human; arginine-rich protein; cancer; inflammation;

XX KW genetic disorder.

XX OS Homo sapiens.

XX PN US2004053250-A1.

XX PD 18-MAR-2004.

XX PF 21-NOV-2002; 2002US-00302172.

XX PR 05-MAR-2001; 2001US-00799451.

XX PR 05-MAR-2002; 2002WO-US005095.

XX PR 20-AUG-2002; 2002US-00225251.

XX PA (TANG/) TANG Y T.

XX PA (XUEA/) XUE A.

XX PA (DRMA/) DRMANAC R T.

XX PI Tang YT, Xue A, Drmanac RT;

DR WPI: 2004-238579/22.
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX Disclosure; SEQ ID NO 215, 51pp; English.
XX The invention relates to an isolated polynucleotide. The methods and
XX compositions of the present invention are useful for the diagnosis and/or
XX treatment of diseases or conditions associated with aberrant expression
XX or activity of the arginine-rich protein-like polypeptides, such as
XX cancer and inflammation. They can also be used in forensics, gene
XX mapping, identification of mutations responsible for genetic disorders,
XX and in assessing biodiversity. The present sequence represents a novel
XX human arginine-rich protein cDNA.
SQ Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;

Query Match 47.3%; Score 1372.2; DB 12; Length 2501;
Best Local Similarity 81.6%; Pred. No. 2,7e-293;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGTGCCCCAGCCACTCCACCTGCGGCTCCCGGACAGTGAAGTGTGCTTAC 141
DB 65 CTATTGATTCCTCCGCGCCCTTGTCTACCTCTGCTGCGCATGAGTGTGCTGTTTCG 124
QY 142 TCCAGCCGCCGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGAGAGCGCCCGGCGC 201
DB 125 GGGGGGCTCCGGGCCACTCCCTCGGCTCCGAGAGCTA-----GCCCGGCGC 172
QY 202 TGGCGGAGCGGCTCATCAATGCTGCTTAACTTGTATGAAGAGAGCGGTGAAGCGC 261
DB 173 TGGCGGAGGCTGATCAATGCTGCTTAACTTGTATGAAGAGAGCGGTGAAGCGC 232
QY 262 ACCATCAAAACAACCTGCGGACCGCTTAAAGTTCCTGAGAGCGCTGAGAGGCA 321
DB 233 ACCATCAAAACAACCTGCGGACCGCTTAAAGTTCCTGAGAGCGCTGAGAGGCA 292
QY 322 CTTACGGGAGTGAAGAGGCAAGAGAGCTCGGGGCTGTGTGAGTGTGAGTGTGAGTGTGAG 381
DB 293 CTTACGGGAGTGAAGAGGCAAGAGAGCTCGGGGCTGTGTGAGTGTGAGTGTGAGTGTGAG 352
QY 382 TCAAGAAAGCAAAATCAAGATGAGCAGGATCTGTGACATACCGAGGAGATTGAGA 441
DB 353 TCCGAAAGAGCAAAATCAAGATGAGCAGGATCTGTGACATACCGAGGAGATTGAGA 412
QY 442 TCAATGTTCACTCAACACCCCACTCACTTTCGATCCATGAGTGTGAGTGTGAGTGTGAG 501
DB 413 TCAATGTTCACTCAACACCCCACTCACTTTCGATCCATGAGTGTGAGTGTGAGTGTGAG 472
QY 502 GCAAGATTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
DB 473 GCAAGATTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532
QY 562 AGCGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
DB 533 AGCGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 592
QY 622 TGCATCTATGCGCACAGAGAGGATGTTTACCGAGATCTCAAGCTGAGAAACATCTTCC 681
DB 593 TGCATCTATGCGCACAGAGAGGATGTTTACCGAGATCTCAAGCTGAGAAACATCTTCC 652
QY 682 TAGATGCAATGAGAAACATCAAGATTGCTGATTTGGGCTCTCAACCTGATCAACAAG 741
DB 653 TAGATGCAATGAGAAACATCAAGATTGCTGATTTGGGCTCTCAACCTGATCAACAAG 712
QY 742 GCAAGTTCCTCAGAGCTTCTGTGGAGCCCTCTTACGCTGCTGAGATGATGATGATGATGATG 801
DB 713 GCAAGTTCCTCAGAGCTTCTGTGGAGCCCTCTTATGCTGCGCAGAGATGATGATGATGATG 772
QY 802 GGAAGCCCTATGTGGGCCAGAGAGTGAAGCTGTGCTGTGGGCTTCTCTGTATCATGCC 861

```

Db      773  GGAAGCCCTACACAGGCGCCAGAGGTGAGCAGCTGCTCTCGGGTTCCTCTCAACATCC 832
Qy      862  TGGTGCATGAGCACCATGCTCCCTTTGACGGGAGAGATCATATAAAACATGCTGAGGAATATCA 921
Db      833  TGGTGCATGAGCACCATGCTCCCTTTGATGAGGATGACATTAAGATCTTGTGTAACAGATCA 892
Qy      922  GTAACGGGAGCTTACCGTGAAGCCGCGCAAGCCGTGCGATGCTGTGAGCTGATTCGGTGC 981
Db      893  GCAACGGGAGCTTACCGGAGCGCACTTAACCCCTGTATGCTGTGAGCTGTGATCCGGTGC 952
Qy      982  TGTTAATGTGTAACCCCAACCCGTGTGAGCCTACCTGAGAGATGTAGCCAGTATGTGTGG 1041
Db      953  TGTTAATGTGTAACCCCAACCCGTGTGAGCCTACCTGAGAGATGTAGCCAGTATGTGTGG 1012
Qy      1042  TCAACTGGGGTTACACACCGGAGTGGGGGACAGAAAGCCCTGCTGAGGGGTGGGACC 1101
Db      1013  TCACCTGGGGCTTACGACCAACCGAGTGGGAGAGAGAGGCTCCGCAATGAGGGTGGGACC 1072
Qy      1102  CTAGTGTGACTTGTGCGCGGAGCTTCATGGCGGACTGTGTTACGTGCTCTCGCGCCCC 1161
Db      1073  CTGGCAGTGACTGTGCGCGGAGCTTCATGGCTGACTGTGCTCGGCTCCCGCCCC 1132
Qy      1162  TCCTGAGAAATGAGCCAGAGTGTGAGCTTCTTCAAGCAGACAGTGC CGGAGGTGAA 1221
Db      1133  TCCTGAGAAATGAGCCAGAGTGTGAGCTTCTTCAAGCAGATGCACTGTGTGGGAA 1192
Qy      1222  GCACTGATGCTGGGCTGGAGCGGAGCAATCTCTTAAGAGTCCGAAAGGAGAAATGCA 1281
Db      1193  GCAACACCCCTGGCTGAGGCGGAGCAATCTCTTAAGAGTCCGAAAGGAGAAATGCA 1252
Qy      1282  TGGCTCAAAATCTGCAAGTGAACCGGCTGAGTACCTCTTCTCGCCCTGCAAGACA 1341
Db      1253  TGGCCAGTGTCTTCCAGATGACAGGCTGTATGACATGCCCATGCCCTGTGGCAAGACA 1312
Qy      1342  GCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCTTCACTGTCAAGGAGGTAC 1401
Db      1313  ACCTCAAGCTGCAAGAGGCAATCTCAAGAAAGAGTGTGAGCTCTGCAAGAGGGTAC 1372
Qy      1402  AGGAGAGCCCTCAGAGACTGAGACCGGTGCTGTATCTCAGGGGACGCTGTCTGTG 1461
Db      1373  AGGAGAGCCCTCAGAGACTGAGACCGGTGCTGTGTGAGCCGAGGAGGCTGCTGCTG 1426
Qy      1462  TATCCCTGCTCCCAAGGAAAGGCAATCTTAAAGAGTCTCGACAGCTGAATCTGTACT 1521
Db      1427  ---CCCTGCTCCCAAGAGGCAATCTTCAAGAGCCCGACAGCGGCAATCTGTACT 1483
Qy      1522  ACTCTCTCCAGAGCCGAGGAGTCTGGGAACTTTAAGCGCCAGTGTGTGTGA 1581
Db      1484  ACTCTCTCCAGAGCCGAGGAGTCTGGGAGTCTTGAAGCGGAGGAGCGTGTGTGA 1543
Qy      1582  GTGGGAGCCCGGTGAGAGAGAGTCTCAAGGCTTCAAGGCTCTCTCTCAAGG 1641
Db      1544  GTGGGAGTCCCAAGAGAGAGAGTCTCGAAGCTTCAAGGCTCTCTCTCATGCAAG 1603
Qy      1642  GCAATCTCAAGTCAATGAGCAAGTCTCCGCAAGGCTTGAAGGCACTAACCCCTAGA 1701
Db      1604  GCAATCTCAAGTCAATGAGCAAGTCTCCGCAAGGCTTGAAGGCTTGAAGGCTTGAAG 1663
Qy      1702  CTTTGGCTCTTGAAGCAAGTGTCTCTCCATCTGAGCCCGGCGGCGGCGGCT 1761
Db      1664  CTTTGGCTCTTGAAGCAAGTGTCTCTCCATCTGAGCCCGGCGGCGGCGGCGGCT 1723
Qy      1762  CAGGGGCTGTGAGTGAAGAGAGAGTCTGTCTCTCGAGTCTTGAAGCAATGTGACTTGC 1821
Db      1724  CAGGGGCTGTGAGTGAAGAGAGAGTCTGTCTCTCGAGTCTTGAAGCAATGTGACTTGC 1783
Qy      1822  CTGAAGCTTCTCCGAAAGCCCACTGAGGGGCTGTGTCTGTGAACAACCTGAGGGGCG 1881
Db      1784  CTGAAGCTTCTCCGAAAGCCCACTGAGGGGCTGTGTCTGTGAACAACCTGAGGGGCG 1843
Qy      1882  TTGAGCAGCTCTCTCAAGAG-----GTCTGAAGCAGTGTGTGAGGAGATCTT 1929

```

```

Db      1844  TTGAGAGAGCCCGCTCAGAGAGGCGCTGGAAGCTGCTGAGAGCGCTGGCGGAGATCTT 1903
Qy      1930  TGGGGATGAGCTGCTTCTTCTGACAGACTGCCAAGAGTGAAGCTGAGCTTACAGACAAG 1999
Db      1904  TGGGGATGAGCTGCTTCTTCTGACAGACTGCCAAGAGTGAAGCTGAGCTTACAGACAAG 1963
Qy      1990  CCTTGAATCTGCTCAAGCTCAGCTGAGAGAGGAGATGTGTCTTGTATGTGGGTAG 2049
Db      1964  CACTGAGGGTCTGCTCAAGCTCAGCTGAGTGAAGTGTGATGTGCTTGTGCTGAG-CCGCTGAG 2022
Qy      2050  GCTTGAAGGGTTTGAAGAGAACCTTGGGTGAGTCTT 2090
Db      2023  GCTTCAAGTCAAGCTGCTTGTGACCCCGAGGGAGATGCTT 2063

```

Search completed: May 15, 2005, 03:21:42
 Job time : 1520.61 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 06:02:36 ; Search time 12445.6 Seconds
(without alignments)
11298.520 Million cell updates/sec

Title: US-09-980-464-4
Perfect score: 2902
Sequence: 1 cactagctgagctccaaagaat.....aaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb ha:*
2: gb htg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	6	AR279568 Sequence
2	2902	100.0	2902	6	AR562226 Sequence
3	2820.4	97.2	2938	10	BC046833 Mus muscu
4	2763.8	92.5	2917	10	BC033302 Mus muscu
5	2102.6	75.2	2900	10	BC081899 Rattus no
6	2066.8	71.2	2929	6	AX380960 Sequence
7	2022.2	69.7	2027	6	AX381046 Sequence
8	1771.4	61.0	2026	6	AX380986 Sequence
9	1593.2	54.9	291762	2	AC125887 Rattus no
10	1379	47.5	3353	6	AX407116 Sequence
11	1375.8	47.4	3395	6	CQ782778 Sequence
12	1375.8	47.4	3395	6	BD127309 Primer fo
13	1375.8	47.4	3395	9	AK074830 Homo sapi
14	1375.8	47.4	3402	9	BC017306 Homo sapi
15	1374.2	47.4	3443	6	AX086949 Sequence
16	1374.2	47.4	3443	6	AX780351 Sequence
17	1370.6	46.9	3360	6	AX642961 Sequence
18	1360.2	46.9	1884	6	AX407118 Sequence
19	1360	46.9	1887	6	CQ714327 Sequence

20	1349.2	46.5	2291	6	AX056394 Sequence
21	1278.6	44.1	252978	2	AC126983 Rattus no
22	1230	42.4	1833	6	AX803413 Sequence
23	1071.4	36.9	1186	6	AX380972 Sequence
24	943.4	32.5	1186	6	AX381029 Sequence
25	849.2	29.3	2616	6	AX430369 Sequence
26	849.2	29.3	2619	6	AX430468 Sequence
27	732	25.2	179414	9	AC098933 Homo sapi
28	728.8	25.1	6389	9	HSMB06181
29	675.4	23.3	741	6	AX380991 Sequence
30	651.8	22.5	1274	5	BX950299 Gallus ga
31	648	22.3	712	6	AX381050 Sequence
32	648	22.3	712	6	AX380990 Sequence
33	644.2	22.2	661	6	AX381052 Sequence
34	643	22.2	734	6	AX381051 Sequence
35	583.2	20.1	588	6	AX381056 Sequence
36	557.4	19.2	660	6	AX380992 Sequence
37	534	18.4	578	6	AX381054 Sequence
38	515	17.7	548	6	AX381053 Sequence
39	513	17.7	521	6	AX381053 Sequence
40	511.6	17.6	579	6	AX380994 Sequence
41	507.6	17.5	6825	6	CQ715220 Sequence
42	507.6	17.5	6828	6	CQ715220 Sequence
43	507.6	17.5	6828	6	CQ715220 Sequence
44	507.6	17.5	6828	9	AB011109 Homo sapi
45	506.4	17.5	1986	6	AX213158 Sequence

ALIGNMENTS

RESULT 1
AR279568
LOCUS AR279568
DEFINITION Sequence 4 from patent US 6514719.
ACCESSION AR279568
VERSION AR279568.1 GI:29714427
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2902)
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.
TITLE Methods for identifying compounds that alter kinase activity
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;
FEATURES
source
1..2902
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2902; DB 6; Length 2902;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2902; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACTAGTGAATCCAAAGATTCCGACAGAGCGTCCGGTGGCTGTGACTGTAG 60
DB 1 CACTAGTGAATCCAAAGATTCCGACAGAGCGTCCGGTGGCTGTGACTGTAG 60
QY 61 CCCGCGGCTCAGCGCGGCTGCTACTGCTGCGCCGACCATCCACTCGCGGTC 120
DB 61 CCCGCGGCTCAGCGCGGCTGCTACTGCTGCGCCGACCATCCACTCGCGGTC 120
QY 121 CCATGAGTGGTGGCTTACTCCAGGCGCCGACGAGCTCCTCGGCTCCGCTGG 180
DB 121 CCATGAGTGGTGGCTTACTCCAGGCGCCGACGAGCTCCTCGGCTCCGCTGG 180
QY 181 CCTCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCCCTAACTGTATGA 240
DB 181 CCTCGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCCCTAACTGTATGA 240
QY 241 AGAAGCAGCGCGGTGAACCGGACCATCAACAACCTGCGGACCGCTACGAGTTCC 300
DB 241 AGAAGCAGCGCGGTGAACCGGACCATCAACAACCTGCGGACCGCTACGAGTTCC 300

Db 241 AGAAGCAGGCGGTGAAGCGGACCATCAAAACAACTGGGGACCCGCTACGAGTTCC 300
Qy 301 TGGAGAGCTGGGCAAGGGCACTTACGGGAAGTGAAGAGGCAAGAGAGCTCGGGC 360
Db 301 TGGAGAGCTGGGCAAGGGCACTTACGGGAAGTGAAGAGGCAAGAGAGCTCGGGC 360
Qy 361 GTCTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGAGTGAAGAGATCTGGCC 420
Db 361 GTCTGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAGAGTGAAGAGATCTGGCC 420
Qy 421 ACATACGAGGAGATTGAGATCAATGTTCTTCACTCAACCAACCCCACTATTCACATCC 480
Db 421 ACATACGAGGAGATTGAGATCAATGTTCTTCACTCAACCAACCCCACTATTCACATCC 480
Qy 481 ATGAAGTGTGAAGATGACGAAAGATTTGATGATGATGATGATGATGATGATGATG 540
Db 481 ATGAAGTGTGAAGATGACGAAAGATTTGATGATGATGATGATGATGATGATGATG 540
Qy 541 ATCTGATGATTAATCATGATGAGCGGCAAGGCTGAGTGAAGCGGGAAGCGGCAATTTCT 600
Db 541 ATCTGATGATTAATCATGATGAGCGGCAAGGCTGAGTGAAGCGGGAAGCGGCAATTTCT 600
Qy 601 TCCGACAGATCTGTCTGCGCTTGAATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 TCCGACAGATCTGTCTGCGCTTGAATGATGATGATGATGATGATGATGATGATGATG 660
Qy 661 TCAAGCTGGAAGAAATCTCTTGAATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TCAAGCTGGAAGAAATCTCTTGAATGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 TCTCAACTGTATCAACAAAGGCAAGTTCTCCAGAGTTCTGAGGAGCCCTCTACAG 780
Db 721 TCTCAACTGTATCAACAAAGGCAAGTTCTCCAGAGTTCTGAGGAGCCCTCTACAG 780
Qy 781 CTTGCGCTGAGATGATCAAGGGAAGCCCTATGAGGCGCAGAGTGAAGAGTGTCTC 840
Db 781 CTTGCGCTGAGATGATCAAGGGAAGCCCTATGAGGCGCAGAGTGAAGAGTGTCTC 840
Qy 841 TGGGCGTTCTCTGTATCATCTGTGTGATGATGATGATGATGATGATGATGATGATG 900
Db 841 TGGGCGTTCTCTGTATCATCTGTGTGATGATGATGATGATGATGATGATGATGATG 900
Qy 901 AAACACTGTGAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 960
Db 901 AAACACTGTGAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 960
Qy 961 CTTGTGCGCTGATCCGCTGTGTTAAATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 CTTGTGCGCTGATCCGCTGTGTTAAATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 1021 ATGTAGCAATCATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 ATGTAGCAATCATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Qy 1081 CCCGCGTGAAGGAGGAGCAACCTAGTGTGATCTTTGCGCGGCTCCATGCGGAGCTGTGT 1140
Db 1081 CCCGCGTGAAGGAGGAGCAACCTAGTGTGATCTTTGCGCGGCTCCATGCGGAGCTGTGT 1140
Qy 1141 TAGGTGCTCTCGCGCGCCCTCTCTGAGAAATGAGAGCAAGTGTGAGAGCTTTCAAGC 1200
Db 1141 TAGGTGCTCTCGCGCGCCCTCTCTGAGAAATGAGAGCAAGTGTGAGAGCTTTCAAGC 1200
Qy 1201 AGCAGCTGCGGAGAGTGAAGCACTGTACTGTGGGCTGAGCGGCAATTTCTTTAAGA 1260
Db 1201 AGCAGCTGCGGAGAGTGAAGCACTGTACTGTGGGCTGAGCGGCAATTTCTTTAAGA 1260
Qy 1261 AGTCCGGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1261 AGTCCGGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy 1321 CTTCTGCGCTTGGCAAGAGAGCTTTAAGCTTCGAAAGGCAATTTCTGAAGAAAGTCT 1380
Db 1321 CTTCTGCGCTTGGCAAGAGAGCTTTAAGCTTCGAAAGGCAATTTCTGAAGAAAGTCT 1380

Qy 1381 CTACCTGTGATGGGAGGTACAGAGAGACCTTCAGAACTTCAGACCGGCTGTATATCTC 1440
Db 1381 CTACCTGTGATGGGAGGTACAGAGAGACCTTCAGAACTTCAGACCGGCTGTATATCTC 1440
Qy 1441 CAGGCGAGCCGTGCTGTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500
Db 1441 CAGGCGAGCCGTGCTGTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1500
Qy 1501 GACAGCTGATCTGTATTAATCTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAG 1560
Db 1501 GACAGCTGATCTGTATTAATCTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAG 1560
Qy 1561 ACAGCAATGATGTGTTGTGATGAGGAGACCCGCTGAGAGAGAGTCTCAAGGCTTCAAG 1620
Db 1561 ACAGCAATGATGTGTTGTGATGAGGAGACCCGCTGAGAGAGAGTCTCAAGGCTTCAAG 1620
Qy 1621 GGGCTCCCTCCACCGGAGAGGAGATTCATCAATGAGAGAGTCTCCGCAAGACCT 1680
Db 1621 GGGCTCCCTCCACCGGAGAGGAGATTCATCAATGAGAGAGTCTCCGCAAGACCT 1680
Qy 1681 TAGAGGCACTAACCCCTAGACCTTTGCTCCCTGAGCAACTGAGCTCTCCATCTCTG 1740
Db 1681 TAGAGGCACTAACCCCTAGACCTTTGCTCCCTGAGCAACTGAGCTCTCCATCTCTG 1740
Qy 1741 CAGCCGCGCCAGCCGCTCTCAGGGCTGTGATGAGAGAGAGATCTGTCTCGAGT 1800
Db 1741 CAGCCGCGCCAGCCGCTCTCAGGGCTGTGATGAGAGAGAGATCTGTCTCGAGT 1800
Qy 1801 CTTTGAACAAATTTGATGCTGAGAGTGTCTCCGAAAGTCTCCGAAAGGAGGCTGTGT 1860
Db 1801 CTTTGAACAAATTTGATGCTGAGAGTGTCTCCGAAAGTCTCCGAAAGGAGGCTGTGT 1860
Qy 1861 CTGTGACAACTGAGGGGCTTGAAGAGGCTCTCTCAGAGGCTTGAAGAGGATGTGTG 1920
Db 1861 CTGTGACAACTGAGGGGCTTGAAGAGGCTCTCTCAGAGGCTTGAAGAGGATGTGTG 1920
Qy 1921 AGGATCTTGGGGGATGAGCTGCTTTTCTCTGACAGCTGCAAGAGTGAATGAGCT 1980
Db 1921 AGGATCTTGGGGGATGAGCTGCTTTTCTCTGACAGCTGCAAGAGTGAATGAGCT 1980
Qy 1981 ACAGCAAGGCTTGAAGAGTGTGCTCAAGAGCTGAGAGAGAGAGAGAGTGTGCTGT 2040
Db 1981 ACAGCAAGGCTTGAAGAGTGTGCTCAAGAGCTGAGAGAGAGAGAGAGTGTGCTGT 2040
Qy 2041 ATGGGATGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGA 2100
Db 2041 ATGGGATGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGA 2100
Qy 2101 GAGTACATCAAGGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGG 2160
Db 2101 GAGTACATCAAGGAGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGG 2160
Qy 2161 ATTGATGTGGAAGAGAGTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGG 2220
Db 2161 ATTGATGTGGAAGAGAGTGAAGAGGCTTGAAGAGGCTTGAAGAGGCTTGAAGAGG 2220
Qy 2221 GGTGCTACCTCTTGTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2280
Db 2221 GGTGCTACCTCTTGTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2280
Qy 2281 CTGACCTTCAATTAAGTTCTGTGTTCAATCAACAGAGGTTAAGACCTTGAATCTCT 2340
Db 2281 CTGACCTTCAATTAAGTTCTGTGTTCAATCAACAGAGGTTAAGACCTTGAATCTCT 2340
Qy 2341 GGAAGGTAATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Db 2341 GGAAGGTAATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Qy 2401 CTGCTGTGCTCTCAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 2460
Db 2401 CTGCTGTGCTCTCAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 2460

Db 1321 |||||CTTCTGCGCCCTGGCAAGGAGCGCTTAAAGTCTCGAAAGGCACTTCTCAAGAAAAAGTCTC 1380
 Qy 1381 CTACCTGTGACGGGAGGATACAGAGAGACCTTCAGGAACTCAGACCCTGCTGATATCTC 1440
 Db 1381 CTACCTGTGACGGGAGGATACAGAGAGACCTTCAGGAACTCAGACCCTGCTGATATCTC 1440
 Qy 1441 CAGGAGACCTGTGCTGCTGATATCCCTGCTCCCAAGGAAAGGATCCTTAAAGAGTCTC 1500
 Db 1441 CAGGAGACCTGTGCTGCTGATATCCCTGCTCCCAAGGAAAGGATCCTTAAAGAGTCTC 1500
 Qy 1501 GACAGCGTGAATCTGTTACTACTCTCTCAGAGCCAGCGAGCTGCGGAACTCTTAG 1560
 Db 1501 GACAGCGTGAATCTGTTACTACTCTCTCAGAGCCAGCGAGCTGCGGAACTCTTAG 1560
 Qy 1561 ACCGCAATGATGTGTTTGTGATGGGAGACCCCTGAGAGCAGAGTCTTCACAGGCTTCAG 1620
 Db 1561 ACCGCAATGATGTGTTTGTGATGGGAGACCCCTGAGAGCAGAGTCTTCACAGGCTTCAG 1620
 Qy 1621 GGCTCCTCTCCACCGGAGGAGGATTCCTCAAACTCAATGGCAATTCCTCCCGCAGAGCT 1680
 Db 1621 GGCTCCTCTCCACCGGAGGAGGATTCCTCAAACTCAATGGCAATTCCTCCCGCAGAGCT 1680
 Qy 1681 TAGAAGGCACTACCCCTAGACCTTTGAGCTCCCTGAGCACTAGGCTCTCCATCTG 1740
 Db 1681 TAGAAGGCACTACCCCTAGACCTTTGAGCTCCCTGAGCACTAGGCTCTCCATCTG 1740
 Qy 1741 CAGCCCGGCGGAGCGGCGCTCAGGGGCTGTGATGAGAGACAGATCTGTCTCCAGT 1800
 Db 1741 CAGCCCGGCGGAGCGGCGCTCAGGGGCTGTGATGAGAGACAGATCTGTCTCCAGT 1800
 Qy 1801 CTTTGAACCAATTGGAATTGCTGAGCTGCTCCCGGAAACCCCACTAGAGGGGCTGTGTGT 1860
 Db 1801 CTTTGAACCAATTGGAATTGCTGAGCTGCTCCCGGAAACCCCACTAGAGGGGCTGTGTGT 1860
 Qy 1861 CTGTGACCAACTGAGGGGAGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAGCGATGTGAGC 1920
 Db 1861 CTGTGACCAACTGAGGGGAGCTTGAAGCAGCTCCCTCAGAAAGTCTGAAGCGATGTGAGC 1920
 Qy 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGCGCAAGAGTGACTGACGCT 1980
 Db 1921 AGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGCGCAAGAGTGACTGACGCT 1980
 Qy 1981 ACAGACAAGCCCTAGAGAACTGCTCAAAAGCTCAGCTAGAGAAAGTGAAGAAATGCG 2040
 Db 1981 ACAGACAAGCCCTAGAGAACTGCTCAAAAGCTCAGCTAGAGAAAGTGAAGAAATGCG 2040
 Qy 2041 ATGGGGTGAAGCTCTGAGAGGGGTTTGAAGAGAACTCCGCTGAGATTCCTCAGATGAATA 2100
 Db 2041 ATGGGGTGAAGCTCTGAGAGGGGTTTGAAGAGAACTCCGCTGAGATTCCTCAGATGAATA 2100
 Qy 2101 GAGTACATCAAGGGGCTCTACGCTGACGCTGACCTGAACCTGAAGATGAAGAAATGCG 2160
 Db 2101 GAGTACATCAAGGGGCTCTACGCTGACGCTGACCTGAACCTGAAGATGAAGAAATGCG 2160
 Qy 2161 ATTGAATGTGAAGAAAGTGAAGAACTCTGCTGCCCGAGTGTATAGTGGGGTGGCTGAA 2220
 Db 2161 ATTGAATGTGAAGAAAGTGAAGAACTCTGCTGCCCGAGTGTATAGTGGGGTGGCTGAA 2220
 Qy 2221 GGTGCTCACTCTCTTGTGTCATAGTGTACCCATGATTCCTCCAGCCCTGTTCTGTGG 2280
 Db 2221 GGTGCTCACTCTCTTGTGTCATAGTGTACCCATGATTCCTCCAGCCCTGTTCTGTGG 2280
 Qy 2281 CTGCACTTCACTAATAGTTTCTGTTTTCATCAACCAACAGGGTGAAGAACCTGACTTCT 2340
 Db 2281 CTGCACTTCACTAATAGTTTCTGTTTTCATCAACCAACAGGGTGAAGAACCTGACTTCT 2340
 Qy 2341 GGGAGGTAATGTGTAGTGTGATCTGCAATATTTAGAGAGAAACAGGCTCTGGTTTCCATCT 2400
 Db 2341 GGGAGGTAATGTGTAGTGTGATCTGCAATATTTAGAGAGAAACAGGCTCTGGTTTCCATCT 2400
 Qy 2401 CTGCTGTGTGCACTCTCAAAAGACCTGGGAGAGCTCGGACGCTGTTGACTCATCTCA 2460
 Db 2401 CTGCTGTGTGCACTCTCAAAAGACCTGGGAGAGCTCGGACGCTGTTGACTCATCTCA 2460

Db 2401 CTGCTGTGTGCACTCTCAAAAGACCTGGGAGAGCTCGGACGCTGTTGACTCATCTCA 2460
 Qy 2461 GGGAGCTCAATGTCCTCCCTGAGACCCCATTTTGAATCTCAGAGACTTTGAAGCTGTT 2520
 Db 2461 GGGAGCTCAATGTCCTCCCTGAGACCCCATTTTGAATCTCAGAGACTTTGAAGCTGTT 2520
 Qy 2521 CCTAGTACCCAGATGTGATGATGATGCTGTGTTCTGAGGCAACGGGACCTGAAATGTGC 2580
 Db 2521 CCTAGTACCCAGATGTGATGATGATGCTGTGTTCTGAGGCAACGGGACCTGAAATGTGC 2580
 Qy 2581 TGAATTAATTAATTTTGTGATTCATCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2640
 Db 2581 TGAATTAATTAATTTTGTGATTCATCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2640
 Qy 2641 TTTTGTGTTTGAAGTGAATTTGCTGTTTCAATATGATGATGCTGTGTTGCGGAGAC 2700
 Db 2641 TTTTGTGTTTGAAGTGAATTTGCTGTTTCAATATGATGATGCTGTGTTGCGGAGAC 2700
 Qy 2701 TCCACTGTGCACTGAAGTTATGTAACAGAAAGTATTTGGCAATGATGTCCTCATTC 2760
 Db 2701 TCCACTGTGCACTGAAGTTATGTAACAGAAAGTATTTGGCAATGATGTCCTCATTC 2760
 Qy 2761 AAGGGGGTGGGGGCGCTTTTCAATGATGCTGTCAGACAGCTGTGATGATGATCTCA 2820
 Db 2761 AAGGGGGTGGGGGCGCTTTTCAATGATGCTGTCAGACAGCTGTGATGATGATCTCA 2820
 Qy 2821 GTCCCTTACACCCCAAGGCTGGGCACTCCCTCATCTTATCTTGTGGCCAAAAA 2880
 Db 2821 GTCCCTTACACCCCAAGGCTGGGCACTCCCTCATCTTATCTTGTGGCCAAAAA 2880
 Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
 Db 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
 RESULT 3
 BC046833
 LOCUS
 DEFINITION Mus musculus RIKEN cDNA 1200013b22 gene, mRNA (CDNA clone MGC:61244 IMAGE:6815659), complete cde.
 ACCESSION BC046833
 VERSION BC046833.1 GI:29145027
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 2938)
 AUTHORS Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Bueltow R.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalske U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2938)
 AUTHORS Strausberg R.

TITLE Direct Submission
JOURNAL Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kuceba, T., Lebeck, M., Melo, A., Schaefer, K., Schaefer, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

FEATURES
source
1..2938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:61244 IMAGE:6816569"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NTH BMAP_PMO"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1..2938
/gene="1200013B22Rik"
/note="synonym: Snack"
/db_xref="LocusID:74137"
/db_xref="MGI:1921387"
81..2000
/gene="1200013B22Rik"
/codon_start=1
/product="1200013B22Rik protein"
/protein_id="AAH4683.1"
/db_xref="GI:29145028"
/db_xref="LocusID:74137"
/db_xref="MGI:1921387"
/translation="MESVALIQRPSPASALASESAPLADGLIKSPKLMKKOYKRHHKHNLRHRYEFTLGKGTGKVKKASGRLVAIKIKDKIKRBDILHIREIIMSLNHPHIIAHEVGRSLVTVFENSRTIVMEYASGDLVDYDISEPRISERDARHPFQIVSALHYCHONGIVARDKLENIILLDANGTKIADFGLSNIYHKKFLQTFCSPLVAPYBEIVNGKPYGPEVDVNSLGLVLIYIVHGTMPDQDHLKVLQISNGAYREPKSPDACGLIRMLAMVNPTRATLEDAVSHMMVNGYTGEOBALBEGHPSGDFGRASMAADMLRBSRPLFENGAVCSFKOHVPGSGTVGLRSHSIXKSRENDMAONICGDPARTSSRPGKSLKPKGIKKKSSTSSGEGVQDPRQIRVPTPQCPVPAVSLPKRGIKKSRORSYSSBPSBEGBLDASVFSGDBVEKSPQASGLILHRKGIILKNGFSRTALEGTPSTFGSLDQLASSHPARPSRPGASVEDSILSESPQDLPERLPTPLRGCVSVNLRGLLEQPPBEGIKRWQBSIGDSFSLTDCQEVTAAYRQALGICYSKLS"
misc_feature
249..1025
/gene="1200013B22Rik"
/note="S_TKC; Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"
/db_xref="CDD:cd00180"

Query Match 97.2%; Score 2820.4; DB 10; Length 2938;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2855; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

43 GCGGCTGACCTCTGAGCCGCGGCTCAGGCGCGCTGCTACTGCTGCGCCGACCCACTTC 102
|||||
Db 1 GCGGCTGACCTCTGAGCCGCGGCTCAGGCGCGCTGCTACTGCTGCGCCGACCCACTTC 60
103 CACCTCGGCGGCCCGGACCATGAGTCCGAGGCTTACTCAGCGCCGACGAGCTC 162
61 CACCTCGGCGGCCCGGACCATGAGTCCGAGGCTTACTCAGCGCCGACGAGCTC 120
163 CCTCGGCTCGCCGCTGCGCTCAGAGCGCCGCGCTGCGGACCGGCTCATCAAGT 222
121 CCTCGGCTCGCCGCTGCGCTCAGAGCGCCGCGCTGCGGACCGGCTCATCAAGT 180
223 CGCTTAACTCTGATGAGAGAGAGCGGCTGAGAGCGGCACATCAAAACAACCTGC 282
181 CGCTTAACTCTGATGAGAGAGAGCGGCTGAGAGCGGCACATCAAAACAACCTGC 240
283 GGCACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGCACTAGGGAGGTGAAGAG 342
241 GGCACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGCACTAGGGAGGTGAAGAG 300
343 CACGAGAGAGCTCGGCGCTGCTGGTGCATCAAGTCCATCAGGAAAGACAAATCAAAG 402
301 CACGAGAGAGCTCGGCGCTGCTGGTGCATCAAGTCCATCAGGAAAGACAAATCAAAG 360
403 ATGAGCAGAGTCTGCTGACATACAGAGGAGATTGAGATCATGTCTTCAACAACC 462
361 ATGAGCAGAGTCTGCTGACATACAGAGGAGATTGAGATCATGTCTTCAACAACC 420
463 CCCACATCATGTCATCATCAATA-----GTGTTGAGATA 498
421 CCCACATCATGTCATCATCAATAAGTGGCAGATCAAGTGTGATGATGTTGAGATA 480
499 GCAGCAGAGTGTGATGTCATGAGATGACGCGGAGGAGTCTGATGATTTATCATCA 558
481 GCAGCAGAGTGTGATGTCATGAGATGACGCGGAGGAGTCTGATGATTTATCATCA 540
559 GTGAGCGGCAACGCGCTGAGTGAAGCGGAGCGGCAAGGATTTCTTCGACAGATCGTGTG 618
541 GTGAGCGGCAACGCGCTGAGTGAAGCGGAGCGGCAAGGATTTCTTCGACAGATCGTGTG 600
619 CCGTGCATCTGCGCACCAAGACGGGATCGTTACCGAGATCTCAAGTGTGAAAACATCC 678
601 CCGTGCATCTGCGCACCAAGACGGGATCGTTACCGAGATCTCAAGTGTGAAAACATCC 660
679 TTCTAGATGCCAATGAGAAACATCAAGATTGCTGCTTGGCTGCCAAGCTGTACACA 738
661 TTCTAGATGCCAATGAGAAACATCAAGATTGCTGCTTGGCTGCCAAGCTGTACACA 720
739 AAGCAAGTTCCTCAGACGTTCTGTGGAGACCTCTCTACGCTCGCTGAGATGATCA 798
721 AAGCAAGTTCCTCAGACGTTCTGTGGAGACCTCTCTACGCTCGCTGAGATGATCA 780
799 ACGGAAAGCCTTATGTGGGCCCAAGAGTGAACAGCTGCTCTGGGGCTTCTCTGTACA 858
781 ACGGAAAGCCTTATGTGGGCCCAAGAGTGAACAGCTGCTCTGGGGCTTCTCTGTACA 840
859 TCCGTGGTGCATGAGACCATGCCCTTTCAGCGGCGAGATCAAAACACTGTGAAAGCAA 918
841 TCCGTGGTGCATGAGACCATGCCCTTTCAGCGGCGAGATCAAAACACTGTGAAAGCAA 900
919 TCAGTAAACGAGGCTTACCGTGAAGCGCCCAAGCGCTCGAGTCCGTGATCCGGT 978
901 TCAGTAAACGAGGCTTACCGTGAAGCGCCCAAGCGCTCGAGTCCGTGATCCGGT 960
979 GGCCTTTAATGTGAACCCCAACCGCTGGGCTCACTGAGAGATGAGCATGATGGT 1038
961 GGCCTTTAATGTGAACCCCAACCGCTGGGCTCACTGAGAGATGAGCATGATGGT 1020

QY 1039 GGGTCACTGGGGTTTACCAACCGAGTCCGGGGAACAGAAACCTTGCTGAGGGTGGC 1098
DB 1021 GGGTCACTGGGGTTTACCAACCGAGTCCGGGGAACAGAAACCTTGCTGAGGGTGGC 1080
QY 1099 ACCCTAGTGTGACTTTGGCCGGGCTCATGCGGAGTGTGTAGTGTCTCTCGGCC 1158
DB 1081 ACCCTAGTGTGACTTTGGCCGGGCTCATGCGGAGTGTGTAGTGTCTCTCGGCC 1140
QY 1159 CCTCTGAGAGATGAGAGCCAGAGTGTGACCTTTCAAGCAGACGTCGCGGAGGTG 1218
DB 1141 CCTCTGAGAGATGAGAGCCAGAGTGTGACCTTTCAAGCAGACGTCGCGGAGGTG 1200
QY 1219 GAAGCAGTGAACCTGGGCTGAGAGGAGCAATCTCTTAAGAATCCCGAAGAGATG 1278
DB 1201 GAAGCAGTGAACCTGGGCTGAGAGGAGCAATCTCTTAAGAATCCCGAAGAGATG 1260
QY 1279 ACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGAGTACCTCTTCTGCGCTGGCAGA 1338
DB 1261 ACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGAGTACCTCTTCTGCGCTGGCAGA 1320
QY 1339 GCAGCCTTTAAGCTTCGAAAGGCAATTCTCAAGAAAAGTCTCTACCTCGTCAAGGGAGG 1398
DB 1321 GCAGCCTTTAAGCTTCGAAAGGCAATTCTCAAGAAAAGTCTCTACCTCGTCAAGGGAGG 1380
QY 1399 TACAGAGAGACCTTCAGAACTCAGACCGGCTGATATCTCAGGGGAGCTCTCGTG 1458
DB 1381 TACAGAGAGACCTTCAGAACTCAGACCGGCTGATATCTCAGGGGAGCTCTCGTG 1440
QY 1459 CTGATCTCTGCTCCCAAGAAAGGCAATCTTAAGAAGTCTCAGACGCTGATCTGAT 1518
DB 1441 CTGATCTCTGCTCCCAAGAAAGGCAATCTTAAGAAGTCTCAGACGCTGATCTGAT 1500
QY 1519 ACTATCTCTCTCCAGAGCCCAAGAGTCTGAGGAACTTTAGACGCCAGTATGTGTTG 1578
DB 1501 ACTATCTCTCTCCAGAGCCCAAGAGTCTGAGGAACTTTAGACGCCAGTATGTGTTG 1560
QY 1579 TGAAGTGGGAGACCCGCTGAGAGAGAGTCTCAGACGCTTCAAGGCTCTCTCCACGCA 1638
DB 1561 TGAAGTGGGAGACCCGCTGAGAGAGAGTCTCAGACGCTTCAAGGCTCTCTCCACGCA 1620
QY 1639 AGGGCAATCTCAAACTCAATGAGCAAGTCTCCGCAAGCTTTAGAAAGGCACTACCCCTA 1698
DB 1621 AGGGCAATCTCAAACTCAATGAGCAAGTCTCTCCGCAAGCTTTAGAAAGGCACTACCCCTA 1680
QY 1699 GCACCTTTGGCTCCCTGAGACCAACTGAGCTCTCTCCATCTCAGACCCGCGCCAGCCGC 1758
DB 1681 GCACCTTTGGCTCCCTGAGACCAACTGAGCTCTCTCCATCTCAGACCCGCGCCAGCCGC 1740
QY 1759 CTTCAAGGAGCTGTGAGTGAAGAGCAATCTGTCTCTCCAGTCTCTTGAACCAATTGACT 1818
DB 1741 CTTCAAGGAGCTGTGAGTGAAGAGCAATCTGTCTCTCCAGTCTCTTGAACCAATTGACT 1800
QY 1819 TGCCTGAACGCTCTCCGAAACCCCACTGAGAGGGCTGTGCTGTGGAACAACCTGAGGG 1878
DB 1801 TGCCTGAACGCTCTCTCCGAAACCCCACTGAGAGGGCTGTGCTGTGGAACAACCTGAGGG 1860
QY 1879 GGGTTCGAGAGCTCTCCCTCAGAAAGTCTGAGCGATGGTGGCAGAAATCTTTGGGGGATA 1938
DB 1861 GGGTTCGAGAGCTCTCCCTCAGAAAGTCTGAGCGATGGTGGCAGAAATCTTTGGGGGATA 1920
QY 1939 GCTGCTTTTCTCTGACAGACTGCGCAAGAGTGAATCTGACAGCTTACAGACAGCCCTAGGAA 1998
DB 1921 GCTGCTTTTCTCTGACAGACTGCGCAAGAGTGAATCTGACAGCTTACAGACAGCCCTAGGAA 1980
QY 1999 TCTGCTCAAAAGCTCAGCTGAGAGAGAGAGATGAGCTCTAGTATGAGGGTGGCTCTAGA 2058
DB 1981 TCTGCTCAAAAGCTCAGCTGAGAGAGAGAGATGAGCTCTAGTATGAGGGTGGCTCTAGA 2040
QY 2059 GGGTTTCAGAGAGACCTGGGTGAGTCTCTCAGTGAATAGATCATCAAGGGCTCT 2118
DB 2041 GGGTTTCAGAGAGACCTGGGTGAGTCTCTCAGTGAATAGATCATCAAGGGCTCT 2100
QY 2119 ACCTTCAGAGCTGACTGAACTGAAAGATGAGAAATCCGATTGATGTGGAAGAAAT 2178

DB 2101 ACCTTCAGAGCTGACTGAACTGAAAGATGAGAAATGCAATGATGTGGAAGAAAT 2160
QY 2179 GGGAAACCTTGCTGCTCCGAGTGTATAGTGGGGTGGCTGAAAGTGTCTACTCTTTGT 2238
DB 2161 GGGAAACCTTGCTGCTCCGAGTGTATAGTGGGGTGGCTGAAAGTGTCTACTCTTTGT 2220
QY 2239 GCCATGAGTGTCAACCAATGACATTTCCACCTGTTCTGTGCTGACCTTACATAGT 2298
DB 2221 GCCATGAGTGTCAACCAATGACATTTCCACCTGTTCTGTGCTGACCTTACATAGT 2280
QY 2299 TTCTGTTTCATCAACCAAGAGTTGAAACCTGATCTTCTGGAGTATGTGTGTG 2358
DB 2281 TTCTGTTTCATCAACCAAGAGTTGAAACCTGATCTTCTGGAGTATGTGTGTG 2340
QY 2359 ACTGCAATTTATGAGAGAGAAACAGGCTCTGGTTTCATCTGCTGCTGATCTCA 2418
DB 2241 ACTGCAATTTATGAGAGAGAAACAGGCTCTGGTTTCATCTGCTGCTGATCTCA 2400
QY 2419 AAGACTGGGAGAGCTGAGACCGCTGTTTGAATTCATCTCAAGGAGACAGATGCCCTG 2478
DB 2401 AAGACTGGGAGAGCTGAGACCGCTGTTTGAATTCATCTCAAGGAGACAGATGCCCTG 2460
QY 2479 GACCCATCTTATGATCTCAGAGACTTGAACCTTGAAGCTTTCTAGTATCCAGATGTG 2538
DB 2461 GACCCATCTTATGATCTCAGAGACTTGAACCTTGAAGCTTTCTAGTATCCAGATGTG 2520
QY 2539 ATGATGCTCTGTTTCTCAGAGCAAGGAGCTTGAATGCTGACTTATTTATTTT 2598
DB 2521 ATGATGCTCTGTTTCTCAGAGCAAGGAGCTTGAATGCTGACTTATTTATTTT 2580
QY 2599 GTGATCTCACTCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2658
DB 2581 GTGATCTCACTCTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2640
QY 2659 TTTTGTCTCTTCAATTAATGTAATGCTGTTCTGGGGAACCTCACTGAGCACTGAG 2718
DB 2641 TTTTGTCTCTTCAATTAATGTAATGCTGTTCTGGGGAACCTCACTGAGCACTGAG 2700
QY 2719 TTTATGTACAGAGAAATTTTGGCAATGATGCTCTTATTTCAAGGGGGTGGGGGCTT 2778
DB 2701 TTTATGTACAGAGAAATTTTGGCAATGATGCTCTTATTTCAAGGGGGTGGGGGCTT 2760
QY 2779 TTTCAATGTATGCTTGAAGACTGTCTGATTTGATCTTCAGTCCCTTACACCCAGG 2838
DB 2761 TTTCAATGTATGCTTGAAGACTGTCTGATTTGATCTTCAGTCCCTTACACCCAGG 2820
QY 2839 CTGGCCACCTCTCTCATCTTCACTGTGTGGCAAAAAAAAAAAAAAAAAAAAAA 2898
DB 2821 CTGGCCACCTCTCTCATCTTCACTGTGTGGCAAAAAAAAAAAAAAAAAAAAAA 2880

RESULT 4
BC033302
LOCUS
DEFINITION
MUS MUSCULUS RIKEN CDNA 1200013b22 gene, mRNA (CDNA clone MGC:30598
IMAGE:3663595), complete cds.
ACCESSION
BC033302
VERSION
BC033302.1 GI:23271085
KEYWORDS
MGC.
SOURCE
MUS MUSCULUS (house mouse)
ORGANISM
MUS MUSCULUS
REFERENCE
AUTHORS
1 (bases 1 to 2917)
Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heise,F.,
Ditchens,K., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L.,
Schneitz,T.B., Brownstein,M.J., Ueda,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Db 733 TCTACGCTCGCCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGACACT 792
Qy 835 GGTCTCTGGGCGTTCTCTGTACATCCTGTGATGAGCAACCATGCCCTTTGACGGGACG 894
Db 793 GGTCTCTGGGCGTTCTCTGTACATCCTGTGATGAGCAACCATGCCCTTTGACGGGACG 852
Qy 895 ATCTATAAACACTGTGAAGCAAAATCAAGTAACGGGGCTTACCGGTGAGCGGCCCAAGCCGT 954
Db 853 ATCTATAAACACTGTGAAGCAAAATCAAGTAACGGGGCTTACCGGTGAGCGGCCCAAGCCGT 912
Qy 955 CCGATGCGCTGTGCGCTGATCCGATGCGCTGTATATGTGTGAACCCCAACCGCTGCGGCAAC 1014
Db 913 CCGATGCGCTGTGCGCTGATCCGATGCGCTGTATATGTGTGAACCCCAACCGCTGCGGCAAC 972
Qy 1015 TGGAGATGTAGCCAGTCAATTGTGTGTCAACTGGGGTTTACCAACCGGAGTGGGGAAAC 1074
Db 973 TGGAGATGTAGCCAGTCAATTGTGTGTCAACTGGGGTTTACCAACCGGAGTGGGGAAAC 1032
Qy 1075 AGGAAGCCCTGCGGTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCCCTCCATGTGCGG 1134
Db 1033 AGGAAGCCCTGCGGTGAGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCCCTCCATGTGCGG 1092
Qy 1135 ACTGTTACGTGCGCTCTCGCGCCCTCTGTGAGAAATGAGCCAAAGGTGTGACGCTTCT 1194
Db 1093 ACTGTTACGTGCGCTCTCTCGCGCCCTCTGTGAGAAATGAGCCAAAGGTGTGACGCTTCT 1152
Qy 1195 TCAAGCAGCAGTGTCCGGGAGGTGAAAGCACTGTACCTGGGCTGGAAGCGCAATTCTC 1254
Db 1153 TCAAGCAGCAGTGTCCGGGAGGTGAAAGCACTGTACCTGGGCTGGAAGCGCAATTCTC 1212
Qy 1255 TTTAAGAGTCCCGAAAGAAATGACATGTGCTCAAAATCTGCAAGGTGACCCGGCTAGG 1314
Db 1213 TTTAAGAGTCCCGAAAGAAATGACATGTGCTCAAAATCTGCAAGGTGACCCGGCTAGG 1272
Qy 1315 ATACCTCTTCTGCGCCCTGSCAAAGACGCTTAAGCTTCCGAAAGGCAATTCTCAAGAAA 1374
Db 1273 ATACCTCTTCTGCGCCCTGSCAAAGACGCTTAAGCTTCCGAAAGGCAATTCTCAAGAAA 1332
Qy 1375 AGTCTCTACCTCTGTCAGGGAGGTAGACAGAGAACCTTCAGAGAACTGACCGGTGCTG 1434
Db 1333 AGTCTCTACCTCTTTCAGGGAGGTAGACAGAGAACCTTCAGAGAACTGACCGGTGCTG 1392
Qy 1435 ATACTCCAGGCGACGCTGTCTCTGTATCTCTGTCTCCAGAGAAAGGCACTCTTAAGA 1494
Db 1393 ATACTCCAGGCGACGCTGTCTCTGTATCTCTGTCTCCAGAGAAAGGCACTCTTAAGA 1452
Qy 1495 AGTCTCGACAGCGGAACTGTGTTACTACTCTCTCCAGAGCCGAGGAGGTCTGGGAAAC 1554
Db 1453 AGTCTCGACAGCGGAACTGTGTTACTACTCTCTCCAGAGCCGAGGAGGTCTGGGAAAC 1512
Qy 1555 TCTTGAACGCGCAGTGTGTGTTGTGTGAGTGGGAGCCCGTGGAGCAGAGTCTCCACAGG 1614
Db 1513 TCTTGAACGCGCAGTGTGTGTTGTGTGAGTGGGAGCCCGTGGAGCAGAGTCTCCACAGG 1572
Qy 1615 CTTCAGGGGCTCTCTCTCAACGCGCAAGGCAATTCTCAAACTCAATGCGAAGTTCTCCGCA 1674
Db 1573 CTTCAGGGGCTCTCTCTCAACGCGCAAGGCAATTCTCAAACTCAATGCGAAGTTCTCCGCA 1632
Qy 1675 CAGCCTTAAGAAGGCACTACCCCTTAGCACTTTGGGCTCCCTGGACCACTGGCCCTCCGCC 1734
Db 1633 CAGCCTTAAGAAGGCACTACCCCTTAGCACTTTGGGCTCCCTGGACCACTGGCCCTCCGCC 1692
Qy 1735 ATCTGAGCGCCGCGCCAGCGCCCTCAAGGGGCTGTGAGTGAAGAGACATCTGTCTCT 1794
Db 1693 ATCTGAGCGCCGCGCCAGCGCCCTCAAGGGGCTGTGAGTGAAGAGACATCTGTCTCT 1752
Qy 1795 CCGAGTCTTTTGAACCAATTGACTTGTCTTAAAGTCTTCCGAAACCCCACTGAGGGGCT 1854
Db 1753 CCGAGTCTTTTGAACCAATTGACTTGTCTTAAAGTCTTCCGAAACCCCACTGAGGGGCT 1812
Qy 1855 GTGTGTGTGTGACCACTGAGGGGGCTTAGCGAGCTCTCCCTCGAAGGTCTGAAGGAT 1914

Db 1813 GTGTGTGTGTGACCAACTGAGGGGGCTTAGAGAGCTCCCTCAGAAAGTCTGAAGGAT 1872
Qy 1915 GGTGGCAGGAATCTTTGGGGGATAGCTGTTTCTCTGACAGACTGCCAAGGTGACTG 1974
Db 1873 GGTGGCAGGAATCTTTGGGGGATAGCTGTTTCTCTGACAGACTGCCAAGGTGACTG 1932
Qy 1975 CAGCCTTAAGAAGCCCTTAGGAATCTGTCTCAAAAGCTCACTGAGGAAGGAATGTGTC 2034
Db 1933 CAGCCTTAAGAAGCCCTTAGGAATCTGTCTCAAAAGCTCACTGAGGAAGGAATGTGTC 1992
Qy 2035 CTTAGTATGGGGTATAGGCTCTGAGAGGGGTGTGACAGAGGAAACCTGTGGTCCGATTCCTCAG 2094
Db 1993 CTTAGTATGGGGTATAGGCTCTGAGAGGGGTGTGACAGAGG-ACCCTGTGGTCCGATTCCTCAG 2051
Qy 2095 TGAATAGATACATCAAGGGGCTTACGTCTGACGCTGACTGAACCTGAAGAATGAGAGA 2154
Db 2052 TGAATAGATACATCAAGGGGCTTACATCTGACGCTGAACCTGAAGAATGAGAGA 2111
Qy 2155 AATCGCATGTATGTGAAAGAAATGGGAACCTTGTGCTGCCCGAGTGTATATGTGGGTGG 2214
Db 2112 AATCGCATGTATGTGAAAGAAATGGGAACCTTGTGCTGCCCGAGTGTATATGTGGGTGG 2171
Qy 2215 CTTGAAGGTGCTTACCTCTTTGTGACATGAGTGTCAACCATGACATTTCCACCCGTT 2274
Db 2172 CTTGAAGGTGCTTACCTCTTTGTGACATGAGTGTCAACCATGACATTTCCCGCCTGTT 2231
Qy 2275 CTCTGGTGCACCTTCAATAGATTCTGTGTTTCATCAACCAACAGGGTTAGAACCTGA 2334
Db 2232 CTCTGGTGCACCTTCAATAGATTCTGTGTTTCATCAACCAACAGGGTTAGAACCTGA 2291
Qy 2335 CTTCCTGGGAGGTATATGTATGATGATCTGCATATTTATAGAGAGAAACAGCTCTGTTT 2394
Db 2292 CTTCCTGGGAGGTATATGTATGATGATCTGTATATTTAGAGAGAAACAGCTCTGTTT 2351
Qy 2395 CCAATCTGTGCTGTGTCATCTCAAGAACCTGGGAGAGCTCGGACCGCTGTTGACTCA 2454
Db 2352 CCAATCTGTGCTGTGTCATCTCAAGAACCTGGGAGAGCTCGGACCGCTGTTGACTCA 2411
Qy 2455 TCTCAAGGGGACAGATGCCCCCTGAGACCCCATCTTAGATCTCAGAGACTTGAACCTTGA 2514
Db 2412 TCTCAAGGGGACAGATGCCCCCTGAGACCCCATCTTAGATCTCAGAGACTTGAACCTTGA 2471
Qy 2515 GCTGTTCTTAGTACCCAGATGTGATGATGCTGTGTTCTCAGGCCAACGGGACCTAGA 2574
Db 2472 GCTGTTCTTAGTACCCAGATGTGATGATGCTGTGTTCTCAGGCCAACGGGACCTAGA 2531
Qy 2575 ATGAGCTGACTTATTTATTTTGTGATCTGACCTCTGTGTTTGTGTTTGTGTTT 2634
Db 2532 ATGAGCTGACTTATTTATTTTGTGATCTGACCTCTGTGTTTGTGTTTGTGTTT 2591
Qy 2635 TGTGTTGTTTGTGTTTAAAGTGAAT-TTGCTGCTTCAATAAATGATGATGTGTTCT 2693
Db 2592 TGTGTTGTTTGTGTTTAAAGGAATATGCTGCTCTTCAATAAATGATGATGTGTTT 2651
Qy 2694 GGGGAATCCACTGTGCTGACTGAAGTTTATGTACAGAGAAATTTGGCAATGATGCC 2753
Db 2652 GGGGAATCCCCTGTGCTGACTGAAGTTTATGTACAGAGAAATTTGGCAATGATGCC 2711
Qy 2754 TCTATTCAAGGGGGGTGGGGGCTTTTCAAAAGTATGCTTAGACACTGTCTGANTGA 2813
Db 2712 TTTATTCAAGGGGGGTGGGGGCTTTTCAAAAGTATGTTTAAAGCCCTGTGTGAATGA 2771
Qy 2814 GTCTCAGTCCCTTCAACCCAGGTGAGCAACCTCCCTCATCTTCACTGTGTGCGCAA 2873
Db 2772 GTTTGAGTCCCTTCAACCCAGGTGAGCGTGGGCCCTCCCTCATTTCTTTGTGTGCCAA 2831
Qy 2874 AAAAAAAAAAAAAAAAAAAAAA 2900
Db 2832 AAAAAAAAAAAAAAAAAAAAAA 2858

RESULT 5
BC081899

LOCUS BC081899 2900 bp mRNA linear ROD 02-SEP-2004
DEFINITION Rattus norvegicus cDNA clone MGC:93817 IMAGE:7110518, complete cds.
ACCESSION BC081899
VERSION BC081899.1 GI:51859485
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2900)
Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Uebin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickerson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E.,
Schmechel, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2900)
Director MGC Project.
Submitted (01-SEP-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: <http://image.llnl.gov>
Series: IRAC Plate: 182 Row: a Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 2900
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93817 IMAGE:7110518"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_id="NIH_MGC_235"
/lab_host="DH10B"
/note="Vector: pExpress1"
73. 1965
/codon_start=1
/product="unknown (protein for MGC:93817)"
/protein_id="AAH81899.1"

/db_xref="GI:51859486"
/translation="MESVALHRRGNLAPSAALATBSARPLADRLISPYPLMKQAV
KRHHKGNLHRYEFLLEKGTGYKPKKBSGRVLAJSIRKDKIKOEODLHTR
REIEMSLNPHIATHVEFENSKVIWMEVASRDLVDYSEIRPLRDBRHF
ROIYVALYCHONGIYHRDLKLNILLDAGNLIKADPGLSNYHKGKPIOTPCGSP
YASPELVNRPYRYPYRVDWSLVLYLYLVHGTMPRPGODHKLHYOISSGAYRBP
PSDACGLIRMLNPLRATLIEDVASHNPNVNGYSTRIQSEALRGHPSGDSGA
SMDWLRSSRPILLENGAKVCSFKHVPVGGSTGPELEROHLSKSRKNDMAQTLQ
NDPAEDTSSRPGKNSLCPKGIILKTKASPSGGEVGPOLRYPVSNPQGPVPAIP
PRGILKKSRSRSGYSSPESGESLADAGVFGVSGDEMGOSPASGRHLRKGIL
KNGKFSRTLEGTABSTFESLDIOLASPHPTARASRPSGVSDESLSSRFDLDLP
EIRPEPLRQCVSDVNDLRRLRLEQFPSESLKRWMBESLSDSCPSLTDCEVTAATQALG
ICSKLS"

ORIGIN
Query Match 72.5%; Score 2102.6; DB 10; Length 2900;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2507; Conservative 0; Mismatches 259; Indels 109; Gaps 14;

51 GACCTTGAGCCGCGGCTCAGCGCGCTGCTACTGCTGCCGACCCACTCCTCGC 110
1 GACCTTGAGCTGCGGCTCTCCGCGGCTGCTGCTCTGCTGCCGACCCCTCGC 60
111 GGTCCCGCACCATGAGTGGTGGCTTACTCAGCGCCCGACGAGCTCCTCGGCG 170
61 CGTCCCGCACCATGAGTGGTGGCTTACTCAGCGCGCGGGAACCTGCTCCGCGC 120
171 TCCGCGCTGCGCTCGGAGACCGCCGCGCGCGGACCGGCTCATCAATCGGCTAA 230
121 TCCGCGCTGCGCACGAGACCGCCGCGCGCTGCGGACCGGCTCATCAATCGGCTAA 180
231 CCTCTGATGAAGAAGACGAGCGGTGAAGCGGACCATGACAAACAACTCGGCGAC 290
181 CCTCTGATGAAGAAGACGAGCGGTGAAGCGGACCATGACAAACAACTCGGCGAC 240
291 TACGAGTTCTTGAGAGCGTGGGCAAGGACCTTACGGAAGGTGAAGAGCGAG 350
241 TACGAGTTCTTGAGAGCGTGGGCAAGGACCTTACGGAAGGTGAAGAGCGAG 300
351 AGCTCGGCGGCTGTGGTGGCTTCAATCATGAGAAACAAATCAAGATGACGAG 410
301 AGCTCGGCGGCTGTGGTGGCTTCAATCATGAGAAACAAATCAAGATGACGAG 360
411 GATCTGTGACATATGAGAGGAGATGAGATGATGATCTTCACTCAACACCCCATC 470
361 GATCTGTGACATATGAGAGGAGATGAGATGATGATCTTCACTCAACACCCCATC 420
471 ATTGCGATCATGAAAGTGTGAGAAATGACGAGAATGTGATGTGATGATATGCC 530
421 ATTGCGATCATGAAAGTGTGAGAAATGACGAGAATGTGATGTGATGATATGCC 480
531 AGCGGAGCGCATCTGTATGATTTATCATGATGAGCGGCGACGCGTGAAGCGGAG 590
481 AGCGGAGCGCATCTGTATGATTTATCATGATGAGCGGCGACGCGTGAAGCGGAG 540
591 AGGATTTCTTCCGACGATCGTGTGCTGCGCTGACTACTGACCAAGGAGATGCT 650
541 AGGATTTCTTCCGACGATCGTGTGCTGCGCTGACTACTGACCAAGGAGATGCT 600
651 CACCGAGATCTCAAGCTGGAACAAATCTTCTGATGCTCAATGGAACATCAAGATGCT 710
601 CACCGAGATCTCAAGCTGGAACAAATCTTCTGATGCTCAATGGAACATCAAGATGCT 660
711 GACTTGGCTCTTCAACCTGTATCAACAAAGGAGGATTTCTTCAAGCTTGTGGAG 770
661 GATTTGGCTCTTCAACCTGTATCAACAAAGGAGGATTTCTTCAAGCTTGTGGAG 720
771 CCTCTCAAGCTCGCTGAGATGATGACAGGAGGAGCTATGAGGCGGACGAGAGTGA 830
721 CCTCTCAAGCTCGCTGAGATGATGACAGGAGGAGCTATGAGGCGGACGAGAGTGA 780
831 AGCTGATCTGAGGCGTCTCTGTATCATCTGATGATGATGATGATGATGATGATG 890

Db 781 AGCTGCTCTGCGGCTTCTTCTGTACATCTGCTGATGCGACCATGCTCCCTTTGAAGG 840
 QY 891 CAGGATATATAAAGCTGTGAGAGCAATCATGATACGCGGCTTACCTGTAGCGGCCAAG 950
 Db 841 CAGAGTATATAAAGCTGTGAGAGCAATCATGATACGCGGCTTACCTGTAGCGGCCAAG 900
 QY 951 CCGTCCGATGCTGTGAGCTGATCCGGTGGCTGTTATGAGAACCCACCCGCGGAGCC 1010
 Db 901 CCGTCTGATGCTGTGAGCTGATCCGGTGGCTGTTATGAGAACCCACCCGCGGAGCC 960
 QY 1011 ACACTGAGAGATGATGACCATGATTTGGTGGCTCAACTGGGGTTACACACCGAGTGGG 1070
 Db 961 ACTGTGAGAGATGATGACCATGATTTGGTGGCTCAACTGGGGTTACACACCGAGTGGG 1020
 QY 1071 GAACAGGAAGCCCTGCTGAGGTGGGACCTTATGATGATGATCTTTGGCCGAGCTTCATG 1130
 Db 1021 GAACAGGAAGCCCTGCTGAGGTGGGACCTTATGATGATGATCTTTGGCCGAGCTTCATG 1080
 QY 1131 GCGGACTGCTTACGCTGCTCTGCGGCCCTCTCTGAGAAATGGAAGCCAGTGTGAGC 1190
 Db 1081 GCGGACTGCTTACGCTGCTCTCTGCGGCCCTCTCTGAGAAATGGAAGCCAGTGTGAGC 1140
 QY 1191 TTCTTCAAGAGACGCTGCGGGAGTGAAGACACTGTAACCTGGGCTGAGCGGCAACAT 1250
 Db 1141 TTCTTCAAGAGACGCTGCGGGAGTGAAGACGCGGACCGGGGCTGAGCGGCAACAT 1200
 QY 1251 TCTCTTAAGAAATCCGAAAGAGATGACATGAGCTCAAAATCTGCAAGGTGACCCGCT 1310
 Db 1201 TCTCTTAAGAAATCCGCAAGAGATGACATGAGCTCAAAATCTGCAAGAAATGACCCAGCT 1260
 QY 1311 GAGGATACCTCTTCTGCGCCCTGCGAGAGAGAGCTTAAAGCTTCCGAAAGGATTTCTAAG 1370
 Db 1261 GAGGATACCTCTTCTGCGCCCTGCGAGAGAGAGCTTAAAGCTTCCGAAAGGATTTCTAAG 1320
 QY 1371 AAAAAGCTCTTACCTGCTGAGGAGAGTACAGAGAGACCTCTAGAACTCAGACCGGATG 1430
 Db 1321 AAAAAGGCTCTTCTGCTGAGGAGAGTACAGAGAGGCTCTCAGAACTCAGACCGGATG 1380
 QY 1431 CCGTATCTCAAGGAGAGCTGCTGCTGCTGATCTCTGCTCTCCAGAGAAAGGATCTCTT 1490
 Db 1381 TCCAAATACCCAGGAGAGCTGCTGCTGCTGATCTCTGCTCTCCAGAGAAAGGATCTCTT 1440
 QY 1491 AAGAAGCTGAGAGAGGATCTGATCTGATCTCTCTCTGAGGAGGAGGAGGATCTGAGG 1550
 Db 1441 AAGAAGCTGAGAGGATCTGATCTGATCTCTCTCTGAGGAGGAGGAGGATCTGAGG 1500
 QY 1551 GAACTCTTGAAGCGCAGTGAATGTTGTGAGTGGGAGCCCGTGAAGAGAGAAATCTTCA 1610
 Db 1501 GAACTCTTGAAGCGAGGTGAATGTTGTGAGTGGGAGCCCAATGAGAGAGAAATCTTCA 1560
 QY 1611 CAGGCTTCAAGGCTCTCTCTTCAACGCAAGGAGCTTCTCAAATCTCAATGGCAAGTTCTCC 1670
 Db 1561 CAAGCTTCAAGG---CGCTCTCAATCGCAAGGAGCTTCTCAAATCTCAATGGCAAGTTCTCC 1617
 QY 1671 CGCAAGGCTTGAAGGAGCACTACCCCTTGAAGACCTTTGGCTCCCTGAGACCAATGGCTCC 1730
 Db 1618 CGCAAGGCTTGAAGGAGCACTACCCCTTGAAGACCTTTGGCTCCCTGAGACCAATGGCTCC 1677
 QY 1731 TCCCATCTGCAAGCGCGGAGCGGAGCGGCTTCAAGGAGCTGAGTGAAGAGCAGCATCTG 1790
 Db 1678 CTTCAATCTCAAGCGCGGAGCGGAGCGGCTTCAAGGAGCTGAGTGAAGAGCAGCATCTG 1737
 QY 1791 TCTCTCGAAGTCTTTTGAACAATTTGAAGCTTGGCTGAAAGTCTTCCGAAACCCCACTAGG 1850
 Db 1738 TCTCTCGAAGTCTTTTGAACAATTTGAAGCTTGGCTGAAAGTCTTCCGAAACCCCACTAGG 1797
 QY 1851 GCGTGTCTCTGTGAGCAACTGAGGGGCTTGAAGAGCTCTCTCAAGAGTCTGAAG 1910
 Db 1798 GCGTGTCTCTGTGAGCAACTGAGGGGCTTGAAGAGCTCTCTCAAGAGTCTGAAG 1857
 QY 1911 CGATGTGAGAGAACTCTTGGGGATAGCTGCTTTCTCTGACAGACTGCGCAAGAGAGTG 1970
 Db 1858 CGATGTGAGAGAACTCTTGGGGATAGCTGCTTTCTCTGACAGACTGCGCAAGAGAGTG 1917

QY 1971 ACTGCAAGCTTACAGACAGAGCCCTTGAAGATCTGCTCAAAAGCTCAGCTGAGAGAGGAGATG 2030
 Db 1918 ACAGCAAGCTTACAGACAGAGCCCTTGAAGATCTGCTCAAAAGCTCAGCTGAGAGAGGAGCA 1977
 QY 2031 GTGCGCTTATG-ATGGGGTAGGCTCTGAGA-GGGTTTGCAGAGGAAACCTGGAGTCCGATTC 2088
 Db 1978 GTGCGCGATGATGGGGTAGGCTCTGAGAGGGGTTTGCAGAGGAAACCTGGAGTCCGATTC 2037
 QY 2089 CTCGATGAAATGAGTATCATCAAG---GGCTTACGCTGACAGCTGACCTGAACCTGA 2144
 Db 2038 CCCAGGGGAGTATGATGATCATCAAGAACTCTCTGCTGCTTCAAGCTGATGAACCTGA 2097
 QY 2145 AGATGAGAGAAATCGCATTTATGTGAAAGAAATG---GAACCTTGTGCTCCGAG 2198
 Db 2098 GGGTGAAGAAATGAGAGATATGAAAGAGCTGACCTTACAGAGTCTTCACTGCGCAGAA 2157
 QY 2199 TGTATATAGGGGT-----GGCTGAAGGTGCTACCTCTTGTGAGCAATGATGAC 2252
 Db 2158 TGTTCAGCGGGGACACAGAGACCTGAAAGTGCCTACCTCTTATGCT--TGAAGTCTAC 2215
 QY 2253 CCATGACATTTCCACCTGCTGCTGCTGACAC-----TTCACATAGTCTCTGT 2304
 Db 2216 CCAATGATCTCCC-CCCTGCTCTGCGCAGTGTCAAGGATGACCAATAGATTCCTGT 2274
 QY 2305 TTCCATCAACACCAAGGCTTGAACCTGACTTCTGAGAGTATGTGTGATGATGCC 2364
 Db 2275 TCGCATGACACCAAGGCTTGAACCTGACTTCTGAGAGTATGTGTGAGAGAACCTGC 2334
 QY 2365 ATTATTAGAGAGAAACAGGCTCTGATTTCAATCTGCTGCTGAGCATCTGAAGAC 2424
 Db 2335 CTATTTTAAAGAGAAACAGGCTCTGATTTCAATCTGCTGCTGAGCATCTGAAGAC 2394
 QY 2425 TGGAAAGACTCGAGCGCTGTTGACTTCAATCTCAAGGAGACCAATGCCCTGAGCCCC 2484
 Db 2395 TAGAAAGACTC-AACTGCTGTTCACTTCAATCTCAAGGAGAC----- 2435
 QY 2485 ATCTTGAATCTCAAGACTTGAACCTTGAAGCTGTTCTTATGACCAAGATGATGAT 2544
 Db 2436 -----CTCAGAGACCTGAGCCTTGAAGCTGTTCTTGAATTAACCAAGATGATGAT 2486
 QY 2545 -GCTGTTTCTCAGGCAACGAGGACCTGAAGTGTGACTTATTTATTTT----- 2596
 Db 2487 ATGCTGTTTCTCAGGCAACGAGGACCAAGATGTGTGACTTATTTATTTTGTGATTC 2546
 QY 2597 -----TGTGATCTCACTTGTGTTTGTGT 2623
 Db 2547 TCACTTGTGTTCTGTTGCTTTTGTGTTTGTGTTTGTGTTGTTGTTGTTCT 2606
 QY 2624 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGAAGTGAATTTGTGCTTTCAATTAATGTAAT 2683
 Db 2607 TGTGTTTGTGTTTGTGTTTGTGTTTGAAGTGAATTTGTGCTTTTGGTGAATGTAAT 2666
 QY 2684 GCTGTGTTCTGAGGAACTCCACGTGACCTGAAGTTTATGTAACAGAGATTTTGGCA 2743
 Db 2667 GCTGTGTTCTGAGGAAACCACTGTGTCAATGAAGTGTGTAACAGAGATTTTGGCA 2726
 QY 2744 ATGATGCTCTTATTAAGAGGGGCTGTTTCAATGATGATGTTGAGACACTG 2803
 Db 2727 GTGATTCCTTCAAT-----GGGGGTGCTTTTCAAGTATGTTTGAAGCACTG 2778
 QY 2804 TCTGATGATGATCTCAGTCCCTTCAACCCAGAGGCTGCGACCTTCCCTCATCT 2858
 Db 2779 TCTGATGATGATCTCAGTCCCTTCAACCCAGAGGCTGCGACCTTCCCTCATCT 2833

RESULT 6
 AX380960
 LOCUS AX380960 2929 bp DNA linear PAT 18-MAR-2002
 DEFINITION Sequence 1 from Patent WO20212456.
 ACCESSION AX380960
 VERSION AX380960.1 GI:19575800
 KEYWORDS

Source	Organism	Rattus sp. Rattus sp. Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1	Drucker, D.J., Rosen, C.F. and Lefebvre, D.L. Ampk-related serine/threonine kinase, designated smark Patent: WO 0212456-A1 14-FEB-2002; 1149336 ONTARIO INC. (CA)
AUTHORS		
TITLE		
JOURNAL		
FEATURES		
source		location/Qualifiers 1..2929 /organism="Rattus sp." /mol_type="unassigned DNA" /db_xref="taxon:10118"
ORIGIN		
Query Match	71.2%;	Score 2066.8; DB 6; Length 2929;
Best Local Similarity	86.4%;	Pred. No. 0;
Matches 2488;	Conservative 0;	Mismatches 277; Indels 115; Gaps 14
QY	49	GTGACCTCTGAGCCCGCGGCTCAGCGCGCGCTGCTACTGCTGCCGACCACTCCACCTC 108
Db	9	GTGACCTCTGAGCCCTGCGGCTCTCCGCGCGCTGCTGCTGCTGCGCCGACCCCTCCGCTC 68
QY	109	GCGGTCCCCGACCACTGAGATCGGTGGCTTAATCTCAAGCGCCGACCGACGGCTCCTCTGG 168
Db	69	GCGGTCCCCGACCACTGAGATCGGTGGCTTAATCAACGGCGCGGGAACCTGGCTCCTCTGG 128
QY	169	CCTCGCGCCCTGAGCCCTCGGAGACGCGCGCGCTGCGCGACGAGGCTCATCAAGTGCCTA 228
Db	129	CCTCGCGCCCTGAGCCCTCGGAGACGCGCGCGCTGCGCGACCGGCTCATCAAGTGCCTCA 188
QY	229	AACCTCTGATGAAGAGGAGCGCGGTGAAGCGGACCATCAAAACAACCTCGCGCAAC 288
Db	189	AACCTCTGATGAAGAGGAGCGCGGTGAAGCGGACCATCAAAACAACCTGAGGCAAC 248
QY	289	GCTACGATTTCTTGAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAAAGGCAAG 348
Db	249	GCTACGATTTCTTGAGACCTCGGCAAGGCACTACGGGAAGGTGAAGAAAGGCAAG 308
QY	349	AGAGCTCGGGCGCTGTGGTGCATCAAGTGCATCAGGAAAGACAAATCAAGATGAGC 408
Db	309	AGAGCTCGGGAGCGCTGTGGTGCATCAAGTGTATAGAGAAAGACAAATCAAGATGAGC 368
QY	409	AGAGCTGCTGACATACGAGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCA 468
Db	369	AGAGCTGCTGACATACGAGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCA 428
QY	469	TCATTGTCATTCATGAAGTGTGTAAGATAGCAGAAATGTGATTTGTCAATGAGATG 528
Db	429	TCATTGTCATTCATGAAGTGTGTAAGACACAGCAATGATGATTTGTCAATGAGATG 488
QY	529	CCAGCCGAGGCGATCTGTATGATTAATCAAGTGAAGCGGCAAGGCTGATGAGCGGAGC 588
Db	489	CCAGCCGAGGCGATCTGTATGATTAATCAAGTGAAGCGGCAAGGCTGATGAGCGGAGC 548
QY	589	CCAGGCAATTTCTTCGACAGATCGTGTGCTCTGCACTAGTCCACCAAGACGGAGTCG 648
Db	549	CCAGGCAATTTCTTCGACAGATCGTGTGCTCTGCACTAGTCCACCAAGACGGAGTCG 608
QY	649	TTGACCGAGATCTCAAGTGGAAAAATCTTTCTGATGCGCAATGAAAACATCAAGATTG 708
Db	609	TTGACCGGAGATCTCAAGCTGGAACAATCTTTCTGATGCGCAAGTGGCAACATCAAGATTG 668
QY	709	CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTCTCTCCAGACGTTCTGTGGGA 768
Db	669	CTGACTTTGGGCTCTCCAACTGTATCACAAAGGCAAGTCTCTCCAGACGTTCTGTGGGA 728
QY	769	GCCCTCTTACGCTCGCTGAGATATGCAACGGAAGCCCTATGTGGGCCCAGAGGTGG 828
Db	729	GCCCTCTTACGCTCTCATGATATGCAACGGAAGCCCTATGTGGGCCCAGAGGTGG 788

[illegible]

Dp	1866	AACGATGGGTGCAGGAATCTTTGGGGGATAGCCCTTTTCTCTGCAACTGCGAAGAGG	1925S
Oy	1969	TGACTGCAGCCTTACAGAACGCCCTAGAAATCTGCTCAAAGCTCAGTGAGAAAGGAGA	2028S
Dp	1926	TGACAGCAGACCCTACAGACAAGCCTTAGAAATCTGCTCGAAGCTGACGTGAAGAGAGAG	1985S
Oy	2029	TGTCGCCCTTAAT-ATGGGGTATAGCTCTGAGAGGGTTTGCAAGAGAACCTTGCTGGATT	2087S
Dp	1986	CAGTCCCCCAGAGTAATGGGGGTATACTTTTAGAGGGGTTTTGCAGAGGAACCTGGGT-AGATT	2044S
Oy	2088	CCTCCAGTGAATAGAGTACATCAAG-----GGCTCAAGCTTCGACGCTGACGTGAACCTGA	2143S
Dp	2045	CCCCAGGGTTGTATAGATCATCAAGAACTCTCTCTCTCTTCAAGCTTGAATTTGAACCTGG	2104S
Oy	2144	AAGATGAGAGAAATGCATTGATGTGAAAAGGATGGGAACCTTGTGCTGCCGAGTGTTA	2203S
Dp	2105	AGGCTGAGAGAAATAGCAGAGATATGAAAAGGACGTGACCTTACAGAGTGTGACGTGAGATG	2164S
Oy	2204	TATGGGGGT---GGCTGAAGGTGECTTACTCTCTTTGGCCATGATGTCACTCCATGCA	2260S
Dp	2165	TGAGCGGACAGAGACTGAAGATGAGTCACTCTTTATATGC--TGATGTCATCCCATGGCA	2222S
Oy	2261	TTTCCCAACCCGTGTTCTGCGCTGCAAC-----TTGACATAGATTCTGTTCATCA	2312S
Dp	2223	TCTCCC-CCCTGCTCTCTGCAAGTGTACAGGGGTATCCACATAGATCTCTGTGCAATCG	2281S
Oy	2313	ACCACCAAGGGTTAGAACCCCTGACTTCTGGAGAGTATGTATGTATGATGATCAATTAATTA	2372S
Dp	2282	ACCACCAAGGGTTAGAACCCCTGACTTCTGGAGAGTATGTATGTATGATGATCAATTAATTA	2341S
Oy	2373	GAGAGGAAACAGCCTCTGTGTTCCATCTCTGCTGTGTGCATCTTCAAGAACCTTGGAGAA	2432S
Dp	2342	AAGAGGAAACAGCCTCTGTGTTCCATCTCTGCTGTGTGCATCTTCAAGAACCTTGGAGAA	2401S
Oy	2433	CTCGGACCGCGTTTGACTTCATCTCAAGGGGACCAAGATGACCCTGAGCCCCATCTTAGA	2492S
Dp	2402	CTC-AACTGCTGTTTCACTTCACTTCAAGGGAC-----	2434S
Oy	2493	TCTCAGAGACTTGAACCTTGAAGCTGTCTTGAAGTACCAGATGTGATGAT-GOTCTGT	2551S
Dp	2435	-CTGAGAGACCTGAGCCTTGAAGCGTGTCTGTATACCAAGACTATGATGATGTCTGT	2493S
Oy	2552	TTTCTCAGGCCAACGGGACCTTGAATGTGTCTGACTTAATTTTGTGATTTCTCACTT	2611S
Dp	2494	TTTCTCAGGCCAACGGGACCTTGAATGTGTCTGACTTAATTTTGTGATTTCTCACTT	2551S
Oy	2612	CTGTTTTTGGTT-----	2624S
Dp	2552	CTGGTTCTGTTCGTTTTTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2611S
Oy	2625	-----TTGT	2678S
Dp	2612	TGTTTGT	2671S
Oy	2679	TGAATGCTGTGTTCTGGGGAACCTCACCTGTGSCCATGGAATTATATGACAGAGAATTT	2738S
Dp	2672	TGAATGCTGTGTTCTGGGGAAGCCACTGTGTCAATTGAAGTGTGTGTGTGTGTGTGTGTGT	2731S
Oy	2739	TGGCAATGATGTCCCTCAATTCAGAGGGGGGTGGGGGCGTTTTTCAATGTATGTCTTGAG	2798S
Dp	2732	TGGCAATGATGTCCCTCAAT-----GGGGGGGTGGCGTTTTTCAAGTATGTCTTGAG	2783S
Oy	2799	CAGTGTCTGATTTGAGTCTTCAGTTCCTTCAACCCAGAGGCTGGGCAACCTCCCTCAATCT	2858S
Dp	2784	CAGTGTCTGATTTGGGTCTCTGTTCCTTCAACCAAGAGGCTGTGCACCTTCCCTCAATCT	2843S

VERSION	AX381046.1	GI:19575868
KEYWORDS	.	
SOURCE	Mus sp.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	Drucker,D.J., Rosen,C.F. and lefebvre,D.L. Ampk-related serine/threonine kinase, designated snark Patent: WO 0212456-A 87 14-FEB-2002; 1149336 ONTARIO INC. (CA)	
JOURNAL		
FEATURES	Location/Qualifiers	
source	1..2027	
ORIGIN	/organism="Mus sp." /mol_type="unassigned DNA" /db_xref="taxon:10095"	
Query Match	69.7%; Score 2022.2;	DB 6;
Best Local Similarity	99.9%; Pred. No. 0;	Length 2027;
Matches 2024;	Conservative 0;	Mismatches 3;
		Indels 0;
		Gaps 0;
QY	49 GTGACCTGTAGCCCGGGGCTCAGCGCGGTGTACTGTCGCCGACCACTCACCTC	108
Db	1 GTGACTTGAGCCCGCGGCTCAGCGCGGTGTACTGTCGCCGACCACTCACCTC	60
QY	109 GCGGTCCCCGACCATGAGTGGGTGACTTACTCAGCGCCGAGCAGGCTCCCTCG	168
Db	61 GGCGTCCCCGACCATGAGTGGGTGACTTACTCAGCGCCGAGCAGGCTCCCTCG	120
QY	169 CCTCGGCCTTGCTCGGAGAGCGCCCGGCGCTGCGGAGCGGCTCATCAAGTCGCTTA	228
Db	121 CCTCGGCCTTGCTCGGAGAGCGCCCGGCGCTGCGGAGCGGCTCATCAAGTCGCTTA	180
QY	229 AACCTCGATGAAGAAGAGCGGTGAAAGCGGACCATCAACAACACTCGGGGACCC	288
Db	181 AACTCTGATGAAGAAGAGCGGTGAAAGCGGACCATCAACAACACTCGGGGACCC	240
QY	289 GCTACGAGTTCTTGAGAGCGCTGGGCAAGGGCACTTACGGGAAGTGAAGAAGGCAAG	348
Db	241 GCTACGAGTTCTTGAGAGCGCTGGGCAAGGGCACTTACGGGAAGTGAAGAAGGCAAG	300
QY	349 AGAGCTGGGGGCGTCTGGTGGCCATCAAGTCATCAGAAAAGCAAATCAAGATGAGC	408
Db	301 AGAGCTGGGGGCGTCTGGTGGCCATCAAGTCATCAGAAAAGCAAATCAAGATGAGC	360
QY	409 AGGATCTGCTGCAATACGAGGGGAGATTGAGATCATGTCTTCACTAACCACTCCCA	468
Db	361 AGGATCTGCTGCAATACGAGGGGAGATTGAGATCATGTCTTCACTAACCACTCCCA	420
QY	469 TCATTGCCATCATGAAGTGTTTAGAAATAGCAGCAAGATTGTATGTCTGAGATG	528
Db	421 TCATTGCCATCATGAAGTGTTTAGAAATAGCAGCAAGATTGTATGTCTGAGATG	480
QY	529 CCAGCCGAGGGGATCTGTATGATTATTAATCAATGATGCGGCAAGCTGATGATGAGCGG	588
Db	481 CCAGCCGAGGGGATCTGTATGATTATTAATCAATGATGCGGCAAGCTGATGATGAGCGG	540
QY	589 CCAAGGCAATTTTCTCCGACAGATGTGTCTGCTGCACTACTGCAACAGAAAGGATTCG	648
Db	541 CCAAGGCAATTTTCTCCGACAGATGTGTCTGCTGCACTACTGCAACAGAAAGGATTCG	600
QY	649 TTGACCGAGATCTCAAGCTGGAAAAACATCCCTTAGATGGCAATGGAAACATCAAGATTG	708
Db	601 TTGACCGAGATCTCAAGCTGGAAAAACATCCCTTAGATGGCAATGGAAACATCAAGATTG	660
QY	709 CTGACTTTGGCTCTCCAACCTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGGGA	768
Db	661 CTGACTTTGGCTCTCCAACCTGTACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGGGA	720
QY	769 GCCCTCTTCAAGCTGCGCTGAGATGATCAACGGAAAGCCCTTATGTGGGCCCAAGGTGG	828
Db	721 GCCCTCTTCAAGCTGCGCTGAGATGATCAACGGAAAGCCCTTATGTGGGCCCAAGGTGG	780

QY	829	ACAGCTGATCTCTGGGCGCTTCTCCGTATCATCTCTGATGACAGCAATGACCTTTGACG	888
Db	781	ACAGCTGATCTCTGGGCGCTTCTCCGTATCATCTCTGATGACAGCAATGACCTTTGACG	840
QY	889	GGCAGGATCATTTAAACACTGTGTGAAGCAAACTAGTAAAGGGGCTTACCGTAGCGGCCA	948
Db	841	GGCAGGATCATTTAAACACTGTGTGAAGCAAACTAGTAAAGGGGCTTACCGTAGCGGCCA	900
QY	949	AGCGCTCCGATGCTGTGGCCTGATCCGGTGGCTGTTATAGTGAACCCACCCGTGGG	1008
Db	901	AGCGCTCCGATGCTGTGGCCTGATCCGGTGGCTGTTATAGTGAACCCACCCGTGGG	960
QY	1009	CCACACTGGAAGATTTAGCCAGTCACTTGTGGGTCAATCTGGGGTTACACACCGAAGTCG	1068
Db	961	CCACACTGGAAGATTTAGCCAGTCACTTGTGGGTCAATCTGGGGTTACACACCGAAGTCG	1020
QY	1069	GGGAACAGAAAGCCCTGCGTGAAGGGTGGGACCCATAGTGTGACTTTTGGCCGGGCTTCA	1128
Db	1021	GGGAACAGAAAGCCCTGCGTGAAGGGTGGGACCCATAGTGTGACTTTTGGCCGGGCTTCA	1080
QY	1129	TGGCGGACTGTGTTACGTGCTCTTCGCGCCCTCTCTGAGAAATGAAACCAAGGTGTGCA	1188
Db	1081	TGGCGGACTGTGTTACGTGCTCTTCGCGCCCTCTCTGAGAAATGAAACCAAGGTGTGCA	1140
QY	1189	GCTTCTTTAAGCAGCAGCTGCCGGGAGGTGAAGCATGTATCTTGGGCTGGAAGCGGCAAC	1248
Db	1141	GCTTCTTTAAGCAGCAGCTGCCGGGAGGTGAAGCATGTATCTTGGGCTGGAAGCGGCAAC	1200
QY	1249	ATTCTCTTAAAGATGCCGAAAGGAATGACATGGCTCAAAATCTGAAGTGAACCGG	1308
Db	1201	ATTCTCTTAAAGATGCCGAAAGGAATGACATGGCTCAAAATCTGAAGTGAACCGG	1260
QY	1309	CTGAGGATATACCTCTCTCGCCCTGGCAAGAGAGGCTTAAAGCTTCCGAAAGGACATTC	1368
Db	1261	CTGAGGATATACCTCTCTCGCCCTGGCAAGAGAGGCTTAAAGCTTCCGAAAGGACATTC	1320
QY	1369	AGAAAAGTCTCTATCACTGTCAAGGGAGGTACAGAGAGAACCTCAGGAATCAGACCGG	1428
Db	1321	AGAAAAGTCTCTATCACTGTGTAGGGAGGTACAGAGAGAACCTCAGGAATCAGACCGG	1380
QY	1429	TGCTGTATATCTCAGGGGACGCTGTCTCTGCTGTATCCCTGTCTCCAGAAAGGACATCC	1488
Db	1381	TGCTGTATATCTCAGGGGACGCTGTCTCTGCTGTATCCCTGTCTCCAGAAAGGACATCC	1440
QY	1489	TTAAGAAGTCTGACAGGCTGATATCTGTATCTACTCCTCTCAAGGCCAGCGAGTCTG	1548
Db	1441	TTAAGAAGTCTGACAGGCTGATATCTGTATCTACTCCTCTCAAGGCCAGCGAGTCTG	1500
QY	1549	GGAATCTTTAGACCCAGTGAATGTGTTGTGTGATGTGGGAGCCCGTGAAGCAGAATCTC	1608
Db	1501	GGAATCTTTAGACCCAGTGAATGTGTTGTGTGATGTGGGAGCCCGTGAAGCAGAATCTC	1560
QY	1609	CACAGGCTTCAAGGGCTCTCTCTTCCACCGCAAGGCAATTCGAACTCATGCGCAAGTCT	1668
Db	1561	CACAGGCTTCAAGGGCTCTCTCTTCCACCGCAAGGCAATTCGAACTCATGCGCAAGTCT	1620
QY	1669	CCCGGACAGCCTTAAAGGCACTAACCTTAGACACTTTGGGTCTCTCTGGAACCACTGGCT	1728
Db	1621	CCCGGACAGCCTTAAAGGCACTAACCTTAGACACTTTGGGTCTCTCTGGAACCACTGGCT	1680
QY	1729	CTTCCCATCTGCAAGCCCGGACCGGACCGCCCTCAAGGGGCTGTGATGAGACAGCATCC	1788
Db	1681	CTTCCCATCTGCAAGCCCGGACCGGACCGCCCTCAAGGGGCTGTGATGAGACAGCATCC	1740
QY	1789	TGTCTCTCGAGTCTTTTGAACCAATTGGAATTGCTGGAACGTCTTCCGAAACCCCATGGA	1848
Db	1741	TGTCTCTCGAGTCTTTTGAACCAATTGGAATTGCTGGAACGTCTTCCGAAACCCCATGGA	1800
QY	1849	GGGGGCTGTGTCTGTGGAACAATGAGGGGCTTGAACAGCTCCCTCAAGAGTCTGGA	1908
Db	1801	GGGGGCTGTGTGTGTGGAACAATGAGGGGCTTGAACAGCTCTCTCTCAAGAGTCTGGA	1860

QY	1909	AGCATGGTGGCAGGAATCCTTGGGGGATACCTGTTTCTCTGACAGACTCCMAAGG	1966
Db	1861	AGCATGGTGGCAGGAATCCTTGGGGGATACCTGTTTCTCTGACAGACTCCMAAGG	1922
QY	1969	TGACTGCAGCTTACAGACAGACCCCTTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGA	2028
Db	1921	TGACTGCAGCTTACAGACAGACCCCTTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGA	1980
QY	2029	TGGTGCCCTTACTATATGGGGTAGGCTCTGAGAGGGTTTGAGAGGAACC	2075
Db	1981	TGGTGCCCTTACTATATGGGGTAGGCTCTGAGAGGGTTTGAGAGGAACC	2027
RESULT 8			
LOCUS	AX380986	2026 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 27 from Patent WO0212456.		
ACCESSION	AX380986		
VERSION	AX380986.1		
KEYWORDS	GI:19575826		
SOURCE			
ORGANISM	Rattus sp.		
	Rattus sp.		
	Eulaliaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1		
AUTHORS	Drucker, D.J., Rosen, C.F. and Iefehyre, D.L.		
TITLE	AMK-related serine/threonine kinase, designated smark		
JOURNAL	Patent: WO 0212456-A 27 14-FEB-2002;		
	1149336 ONTARIO INC. (CA)		
FEATURES	Location/Qualifiers		
source	1..2026		
	/organism="Rattus sp."		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:10118"		
ORIGIN			
Query/Match			
Best Local Similarity 93.1%; Pred. No. 0; Mismatches 136; Indels 5; Gaps 3			
Matches 1888; Conservative 0;			
QY	49	GTGACCTTGAAGCCCGGCTCAGCGCGCTGCTACTCTGCCGACCACTCCACTC	108
Db	1	GTGACCTTGAAGCCCGGCTCAGCGCGCTGCTACTCTGCCGACCACTCCACTC	60
QY	109	GGCGTCCCGGACCATGAGTGGTGGCTTACTCAGCGCCGAGCCAGGCTCCTCGG	168
Db	61	GCGGTCCTCGGACCATGAGTGGTGGCTTACTCAGCGCCGAGCCAGGCTCCTCGG	120
QY	169	CCTCGGCTTGGGCTCGGAGCGCGCGGCTGCGGAGCGGGCTCATCAAGTGCCTA	228
Db	121	CCTCGGCTTGGGCTCGGAGCGCGCGGCTGCGGAGCGGGCTCATCAAGTGCCTA	180
QY	229	AACCTTGATGAAGAGCAGCGGTGAGCGGACCATCAAAACAACTCGGCAAC	288
Db	181	AACCTTGATGAAGAGCAGCGGTGAGCGGACCATCAAAACAACTCGGCAAC	240
QY	289	GCTACGAGTCTCTGAGACGCTGGGCAAGGGCACTACGGGAAGGTGAAGAGGACAG	348
Db	241	GCTACGAGTCTCTGAGACGCTGGGCAAGGGCACTACGGGAAGGTGAAGAGGACAG	300
QY	349	AGAGCTGGGGGCGTCTGGTGGCATCAAGTCACTCAGGAAGAAGAAATCAAGATGAGC	408
Db	301	AGAGCTGGGGGCGCTGGTGGCATCAAGTCTATCAGGAAGAGCAAAATCAAGATGAGC	360
QY	409	AGGATCTGCTGCAACATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCA	468
Db	361	AGGATCTGCTGCAACATACGAGGAGATTTAGATCATGTCTTCACTCAACCAACCCCA	420
QY	469	TCAATTGCATCAAGAAAGTGTTTAGAAATACAGAGAAATGTGATGTCTAGAGATG	528
Db	421	TCAATTGCATCAAGAAAGTGTTTAGAAACACAGAAATGTGATGTCTAGAGATG	480

QY 529 CAGCCGAGGCGATCTGTATGATTAACATCAGTAGAGCGGCGACGGCTGAGTGGAGCGGAGC 588
 DB 481 CAGCCGAGGCGATCTGTATGATTAACATCAGTAGAGCGGCGACGGCTGAGTGGAGCGGAGC 540
 QY 589 CAGGCAATTTCTTCCGACAGATCGTGTCTGCTGCACTACTGCGACCGAGACCGGAGTGC 648
 DB 541 CAGGCAATTTCTTCCGACAGATCGTGTCTGCTGCACTACTGCGACCGAGACCGGAGTGC 600
 QY 649 TTACCGAGATCTCAGAGTGGAAAAACATCCCTTCTAGATGCGAATGGAACATCAAGATTG 708
 DB 601 TTACCGAGATCTCAGAGTGGAAAAACATCCCTTCTAGATGCGAATGGAACATCAAGATTG 660
 QY 709 CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTCTCTCAAGCTTTCTGTGGA 768
 DB 661 CTGACTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTCTCTCAAGCTTTCTGTGGA 720
 QY 769 GCGCTCTTACGCTCGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCGCAGAGTGG 828
 DB 721 GCGCTCTTACGCTCGCTGAGATGATCAACGGGAAAGCCCTATGTGGGCGCAGAGTGG 780
 QY 829 ACAGCTGTCTTGGGGGCTTCTCTGTACATCTCTGTGATGATGCAACATGCTCTTTGACG 888
 DB 781 ACAGCTGTCTTGGGGGCTTCTCTGTACATCTCTGTGATGATGCAACATGCTCTTTGACG 840
 QY 889 GGCAGATCATTAACACTGTGAAAGCAATCAGTAACGGGCGCTTACCGTGAAGCGGCCA 948
 DB 841 GGCAGATCATTAACACTGTGAAAGCAATCAGTAACGGGCGCTTACCGTGAAGCGGCCA 900
 QY 949 AGCGTCCGATCGCTGTGGCTGTATCCGGTGGCTTTAATGTGAAACCCACCGCTGGG 1008
 DB 901 AGCGTCCGATCGCTGTGGCTGTATCCGGTGGCTTTAATGTGAAATCCATCCGTGGG 960
 QY 1009 CCACTGAGAGATGTAGCCAGTATGTGTGTGTCACTGGGGTTTACCAACCGGAGTGC 1068
 DB 961 CCACTGAGAGATGTAGCCAGTATGTGTGTGTCACTGGGGTTTACCAACCGGAGTGC 1020
 QY 1069 GGGAAACAGAAAGCCCTGCGTGAAGGGGCGACCCCTAAGTGGTGAATTTGGCGGCGCTTCA 1128
 DB 1021 GGGAAACAGAAAGCCCTGCGTGAAGGGGCGACCCCTAAGTGGTGAATTTGGCGGCGCTTCA 1080
 QY 1129 TGGCGAATGTTACGTCTGCTCTCGGCGCCCTCTCTGAGATGAGCCAAAGTGTGCA 1188
 DB 1081 TGGCGAATGTTACGTCTGCTCTCGGCGCCCTCTCTGAGATGAGCCAAAGTGTGCA 1140
 QY 1189 GCTTCTTCAAGCAGAGTGTGCGGAGGTGGAACATCTGTCTGTGGCTGTGAGCGGCAAC 1248
 DB 1141 GCTTCTTCAAGCAGAGTGTGCGGAGGTGGAACATCTGTCTGTGGCTGTGAGCGGCAAC 1200
 QY 1249 ATTCTCTTAAGAGATCCCGAAAGGAGATGATGAGCTCAAAATCTGCAAGGTGACCCGG 1308
 DB 1201 ATTCTCTTAAGAGATCCCGCAAGAGATGATGAGCTCAAAATCTGCAAGGTGACCCGG 1260
 QY 1309 CTGAGATACCTCTTCTGCGCCCTGCAAGAGACCTTAAAGCTTCCGAAAGCATTTCTCA 1368
 DB 1261 TTGAAAGATACCTCTTCTGCGCCCTGCAAGAGACCTTAAAGCTTCCGAAAGCATTTCTCA 1320
 QY 1369 AGAAAAAGTCTCTTACTCTGTCAGGGAGGTACAGAGAGACCTTCAAGAACTCAGACCGG 1428
 DB 1321 AGAAAAAGTCTCTTACTCTGTCAGGGAGGTACAGAGAGACCTTCAAGAACTCAGACCGG 1380
 QY 1429 TGCCTGATACCTCAGGGGAGCTGTCTCTGCTGTATCCCTGCTCCAAAGAAAGGATCC 1488
 DB 1381 TGCCTGATACCTCAGGGGAGCTGTCTCTGCTGTATCCCTGCTCCAAAGAAAGGATCC 1440
 QY 1489 TTAAAGATCTCAGCAGCGTGAATCTGTGTTACTACTCTCTCAGAGCCCGAGAGCTG 1548
 DB 1441 TTAAAGATCTCAGCAGCGTGAATCTGTGTTACTACTCTCTCAGAGCCCGAGAGCTG 1500
 QY 1549 GGGAACTCTTAAGACCGCAGTGAATGTTTGTGAGTGGGAGACCCCGTGAAGCAAGTCTC 1608
 DB 1501 GGGAACTCTTAAGACCGCAGTGAATGTTTGTGAGTGGGAGACCCCGTGAAGCAAGTCTC 1560
 QY 1609 CACAGGCTTACGGGCTCTCTCTCCACCGCAAGGCAATTCATCAACTCAATGGCAAGTCT 1668

DB 1561 CACAGCTTACGGG---CGCTTCATGCGAAGGCGATCTCAAACTCAATGGCAAGTTT 1617
 QY 1669 CCGGACAGGCTTTAAGAGGACATACCCCTTAGACCTTTGGCTCTCTGAGCAATGGCT 1728
 DB 1618 CCGGACAGGCTTTAAGAGGACATACCCCTTAGACCTTTGGCTCTCTGAGCAATGGCT 1677
 QY 1729 CTTCCCATCTCTGAGAGCCCGGCGCAGAGCCCGCTCAGGGGCTGTGAGTGAAGACAGCTCC 1788
 DB 1678 CCGCTCATCTTACAGCCCGGCGCAGAGCCCGCTCCTGAGAGCTGTGAGTGAAGACAGCTCC 1737
 QY 1789 TGTCTCCGAGCTCTTGAACCAATTGGAATCTGTGCTGTAACGCTCTTCCGAAACCCCACTGA 1848
 DB 1738 TGTCTCCGAGCTCTTGAACCAATTGGAATCTGTGCTGTAACGCTCTTCCGAAACCCCACTGA 1797
 QY 1849 GGGGCTGTGTGTGTGTGAGCAACCTGAGGGGGCTTGAAGAGCTCTCTCAGAAAGTCTGA 1908
 DB 1798 GGAAGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTGAAGAGCTCTCTCAGAAAGCTTGA 1857
 QY 1909 AGGATGTGTGTGTGTGAGCAATCTTGGGGGATGTGCTTTTCTTGAACAGATCTGCCAAGAG 1968
 DB 1858 AAGATGTGTGTGTGTGAGCAATCTTGGGGGATGTGCTTTTCTTGAACAGATCTGCCAAGAG 1917
 QY 1969 TGACTGAGGCTTACAGCAAGAGCCCTAGAGATCTGTCTCAAGTCTCAGCTGAGAGAGGAGA 2028
 DB 1918 TGAAGAGAGCTTACAGCAAGAGCCCTAGAGATCTGTCTGAAAGCTCAGCTGAGAGAGAGG 1977
 QY 2029 TGGTCCCTAGT-ATGGGATAGGCTCTGAGA-GGGTTTGAGAGAAAC 2075
 DB 1978 CAGTGCCCGCAGTATGAGGGGTAGACTTTAAGAGGGGTTTGAGAGAGAAC 2026

RESULT 9
 AC125887
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
 *** 13 unordered pieces.
 AC125887 291762 bp DNA linear HTG 15-NOV-2002
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 1 (bases 1 to 291762)
 Muzny,D,Marie,, Metzker,M, Lee, S, Abramson, S, Adams, C, Alder, J, Allen, C, Allen, H, Alsbrooks, S, Amin, A, Anguiano, D, Araya-Lebeche, V, Aoyagi, A, Ayodeji, M, Baca, E, Baden, H, Baldwin, D, Bandaranaike, D, Barber, M, Barnstead, M, Benahmed, F, Biswal, K, Blair, J, Blankenburg, K, Blyth, P, Brown, M, Bryant, N, Buhay, C, Burch, P, Burrell, K, Calderon, E, Cardenas, V, Carter, K, Cavazos, I, Caesar, H, Center, A, Chacko, J, Chavez, D, Chen, G, Chen, R, Chen, Y, Chen, Z, Chu, J, Cleveland, C, Cockrell, R, Cox, C, Coyle, M, Cree, A, D'Souza, L, Davila, M, L, Davis, C, Davy-Carroil, L, De Anda, C, Dedrich, D, Delgado, O, Denson, S, Deramo, C, Ding, Y, Dinh, B, Divya, K, Draper, H, Dugan-Rocha, S, Dunn, A, Durbin, K, Duval, B, Eaves, K, Egan, A, Escotto, M, Eugene, C, Evans, C, A, Falls, T, Fan, G, Fernandez, S, Finley, M, Flagg, N, Forbes, L, Foster, M, Foster, P, Frazer, C, M, Gabriel, A, Ganta, R, Garcia, A, Garner, T, Garza, M, Gebregeorgis, E, Geer, K, Gail, R, Grady, M, Guerra, W, Guevara, W, Gunaratne, P, Haaland, W, Hamill, C, Hamilton, C, Hamilton, K, Harvey, Y, Havlak, P, Hawes, A, Henderson, N, Hernandez, J, Hernandez, R, Hines, S, Hladun, S, L, Hodgson, A, Hognes, M, Hollins, B, Howells, S, Huyl, S, Hume, J, Idlebird, D, Jackson, L, Jackson, L, Jacob, L, Jiang, H, Johnson, B, Johnson, R, Jolivet, A, Karpathy, S, Kelly, S, Kelly, S, Khan, Z, King, L, Kovar, C, Kowitz, C, Kraft, C, L, Lebow, H, Levan, J, Lewis, L, Li, Z, Liu, J, Liu, J, Liu, W, Liu, Y, London, P, Longacre, S, Lopez, J, Lorensen, L, Louie, L, Lozano, R, J, Lu, X, Ma, J, Maheshwari, M, Mahindaratne, H, Mahmoud, M, Mallory, K, Mangum, A, Mangum, B, Mapa, P, Martin, K, Martin, R, Martinez, B.

Query Match	Best Local Similarity	Matches 1601; Conservative	54.9%; Score 1593.2; DB 2; Length 291762;
Db 55	TCGAGCCCGGAGCTCAGGCGCGGTGCTACGCTGCGCCGACCACTCAGCTCGCGCTC	114	
Db 274674	TCGAGCCCGGAGCTCAGGCGCGGTGCTACGCTGCGCCGACCACTCAGCTCGCGCTC	114	
Db 115	CCGCAACATGAGAGTGTGAGCTTACTCAACGCGCGAGGCAAGCTCCCTCGGCTCCG	174	
Db 274734	CCGCAACATGAGAGTGTGAGCTTACTCAACGCGCGAGGCAAGCTCCCTCGGCTCCG	174	
Db 175	CCCTGGCTTCGAGAGCCCGCGCGCTGCGGAGCGGCTCATCAAGTCCCTTAACCTC	234	
Db 274794	CCCTGGCTTCGAGAGCCCGCGCGCTGCGGAGCGGCTCATCAAGTCCCTTAACCTC	234	
Db 235	TGATGAGAGAGGAGGCGGTGAAGCGGCAACATCAACAAACACCTGCGGCAACCGCTACG	294	
Db 274854	TGATGAGAGAGGAGGCGGTGAAGCGGCAACATCAACAAACACCTGCGGCAACCGCTACG	294	
Db 295	AGTTCTGAGAGAGCTGAGGCAAGGCAACCTACCGGAAAGTGAAGAGGCAAGAGAGCT	354	
Db 274914	AGTTCTGAGAGAGCTGAGGCAAGGCAACCTACCGGAAAGTGAAGAGGCAAGAGAGCT	354	
Db 355	CGGAGGCTGTGAGGCTCAATGATCAATGAGAGAGAGAGAGAGAGAGAGCT	414	
Db 274974	CGGAGGCTGTGAGGCTCAATGATCAATGAGAGAGAGAGAGAGAGAGAGAGCT	414	
Db 415	TGCTGACATACGAGAGGAGATTGATCATGTCTTCACTCAACACCCCAATCATTTG	474	
Db 275034	TGCTGACATACGAGAGGAGATTGATCATGTCTTCACTCAACACCCCAATCATTTG	474	

Db 644 TGCACTATGGCCATCAGAAAGAGTTGTCCAGGAGATCTCAAGCTGGAGAAATCCTCT 703
 Qy 682 TAGATGCAATGAAACATCAAGATTTGCTGCTTGGCTCTCCAACTGTACCAAG 741
 Db 704 TGGATGCAATGGAATATCAAGATTTGCTGCTTGGCTCTCCAACTGTACCAAG 763
 Qy 742 GCAAGTTCTTCAGACGTTCTGTGGAGCCCTCTCAAGCTCTGCTGAGATGATCAAG 801
 Db 764 GCAAGTTCTTCAGACATTTCTGTGGAGCCCTCTCAAGCTCTGCTGAGATGATCAAG 823
 Qy 802 GGAAGCCCTATATGTTGGGCCCCAGAGTGTGACAGTGTCTGTGGAGTTCTCTGTACATCC 861
 Db 824 GGAAGCCCTACACAGGCCCCAGAGTGTGACAGTGTCTGTGGAGTTCTCTGTACATCC 883
 Qy 862 TGGGCAATGGAACCATGCTTTTGAACGGGAGAGATCAATAAACACTGTGAAAGAAATCA 921
 Db 884 TGGGCAATGGAACCATGCTTTTGAATGGGAGATGACATTAAGATCTTAATGAAACAGATCA 943
 Qy 922 GTAAACGGGGCTTACCGTGAAGCCGCCAAGCCGTCGATGCTGTGGCTGATCCGGTGGC 981
 Db 944 GCAACGGGGCTTACCGGGAAGCCCTGAATGCTGTGGCTGATCCGGTGGC 1003
 Qy 982 TGTATATGTTGAACCCCAACCTGTGGGCGACACTGAGAGATGTAGCCAGTATGGTGGG 1041
 Db 1004 TGTATATGTTGAACCCCAACCTGTGGGCGACACTGAGAGATGTGGCCAGTCACTGTGGG 1063
 Qy 1042 TCAACTGGGGTTACCAACCGGAGTCCGGGGAACAGAAAGCCCTGCTGAGGGTGGGCAAC 1101
 Db 1064 TCAACTGGGGTTACCGCAACCGGAGTGGAGAGAGAGGCTCCGATGAGGGTGGGCAAC 1123
 Qy 1102 CTAGTGTGACTTTGGCCGGGCTTCATATGGCGGACTGGTTACGTGTGCTCTCCGCCCCC 1161
 Db 1124 CTGCAATGACTCTGCGCGGCTTCATATGGCGGACTGGTTACGTGTGCTCTCCGCCCCC 1183
 Qy 1162 TCTCTGAGATGAGCCAGAGTGTGACGTTCTTCAAGCAGACGTCGCGGAGGTGAA 1221
 Db 1184 TCTCTGAGATGAGCCAGAGTGTGACGTTCTTCAAGCAGACGTCGCGGAGGTGAA 1243
 Qy 1222 GCACTGATCTGGGCTGGAAGGCAATTTCTTAAAGATCCCGAAGAGAGATGACA 1281
 Db 1244 GCAACACCCCTGGGCTGGAAGGCAATTTCTTAAAGATCCCGAAGAGAGATGACA 1303
 Qy 1282 TGGCTCAAAATCTGCAAGTGAACCCGGCTGAGATACCTCTTCCGCTGGAAGAGA 1341
 Db 1304 TGGCTCAATCTCTCAAGTGAACCGGCTGATGACATGCGCCATGCGCTTGGCAAGACA 1363
 Qy 1342 GCCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTTCACTCTGCAAGGAGTAC 1401
 Db 1364 ACTCAAGCTGCAAGAGGCAATTTCTCAAGAAAGTGTGACGCTCTGCAAGAGGGTAC 1423
 Qy 1402 AGAGAGACCTGCAAGAACTCAGACCGGTGCTGATATCTCAGGGGAGGCTGTCCCTGCTG 1461
 Db 1424 AGAGAGACCTGCAAGAGGCTCAGGCAATCTCTGGAAGCCAGGGGAGGCTGCTCC----- 1477
 Qy 1462 TATCCGCTGCTCCGAAAGGAGCAATCTTAAAGTCTGCAAGGCTGAATCTGCTTACT 1521
 Db 1478 ---CCCTGCTCCGAAAGGCAATCTTCAAGAAAGCCCGCAAGGCGAGCTGCTGCTACT 1534
 Qy 1522 ACTCTCTCCAGAGCCGAGAGTCTGGGGAATCTTGAAGCGCAGAGTATGTTTGTGA 1581
 Db 1535 ACTCTCTCCGAGCCGAGAGTCTGGGGAATCTTGAAGCGCAGAGGCGAGCTGTTTGTGA 1594
 Qy 1582 GTGGGAGCCCGTGAAGAGATCTCAAGGCTTCAAGGCTCTCTCTCAAGGCAAGG 1641
 Db 1595 GTGGGAGTCCCAAGAGAGAGGCTCCGAAAGTCAAGGCTCTCTCTCATGCAAG 1654
 Qy 1642 GCATTTCAAACTCAATGGAAGTCTCCGCAAGCCTTGAAGAGCACTACCCCTAGCA 1701
 Db 1655 GCATTTCAAACTCAATGGAAGTCTCTCCAGAGCCTTGAAGTCCGCGGCCCAACCA 1714
 Qy 1702 CTTTGGCTCCCTGAGCAACTGAGCTTCTCTCCATCTGCAAGCCCGGCGCAAGCCCTT 1761
 Db 1715 CTTTGGCTCCCTGAGTGAATCTGCGCCCACTGCGCCCTGCGCGGCGCAAGCCCTT 1774

Qy 1762 CAGGGGCTGTGAGTGAAGACAGACTCTGTCTCCAGTCTTTGACCAATGACTTGC 1821
 Db 1775 CAGGGGCTGTGAGGAGAGAGAGATCTGTCTCTGAGTCTTTGACCAAGCTGACTTGC 1834
 Qy 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGACAACTGAGGGGCG 1881
 Db 1835 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGTGTGTGTGTGACAACTGAGGGGCG 1894
 Qy 1882 TTGAGCAGCCCTCCGAGAG-----GTGAGAGCGATGTGGCAGGAATCCT 1929
 Db 1895 TTGAGAGCCCTCCGAGAGGCTCCGAGAGCTGAGAGGCTGTGGCGCAGGAATCCTT 1954
 Qy 1930 TGGGGATGAGCTGCTTTCTGTGACAGACTGCCAAGAGTGTGACTGACGCTTACAGCAAG 1989
 Db 1955 TGGGGAGAGCTGCTTTTCTGTGACAGACTGCCAAGAGTGTGACTGACGCTTACAGCAAG 2014
 Qy 1990 CCTGAAGATCTGCTCAAGAGTCACTGAGAGAGAGAGTGTGCTTATGATGGGTAG 2049
 Db 2015 CACTGAGAGTCTGCTCAAGAGTCACTGAGTGTGAGTGTGAGTGTGCTTCCAG-CCCGGTGAG 2073
 Qy 2050 GCTCTGAGAGGTTTGAAGAGAACCTGGGTGGATTTCTCC 2092
 Db 2074 GCTCTGAGAGTCACTGCTGTGACCCGAGGAGAGATGCTTC 2116

RESULT 11
 LOCUS CQ782778 3395 bp DNA linear PAT 17-MAR-2004
 DEFINITION Sequence 2918 from Patent EP1396543.
 ACCESSION CQ782778
 VERSION CQ782778.1 GI:45502721
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Oca, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
 Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
 Koga, H.
 TITLE Primers for synthesizing full length cDNA clones and their use
 JOURNAL Patent: EP 1396543-A 2918 10-MAR-2004;
 Research Association for Biotechnology (JP)
 FEATURES
 source
 1..3395
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 118..2004
 /note="unamed protein product"
 CDS
 /db_xref="GI:45502722"
 /translation="MESLTFARBSGPTPSAELRLPLAEGILKSPKPLMKKAVRHH
 HKNLPRVRELTIGGTGYKYGKAGESGRLVLAISIRDKTKDODMLHRIIE
 IMSLNPHILAIHVEPNSKIYIVMEYSRDLUYIYSRODLSERBAHFROLY
 SAVHYCHQNEVHRLDLNLILDLANENIKIADGSLNLHQGFLDTFGSPLYASP
 EIVNGKPYTPEVDWSLGLVLYLVHGTGPPFDGDKHLIVKQISNGAYRBPFRSDA
 CGILRWLNLNPTLRATLEDDVAISHWVNWGATATVGOEAPHEGSHGSDSARASMA
 MLRRSRLPLENGAKVCSFKQAHAGGGGSTTGLEROSHLSKSKENDMAOHLSDTA
 DPTAHPKSNLKLPGIILKKYVASAEVGOEDPELSPIPASGGOAAPLLPKGIILK
 KPRORESGYISSPESRSGELLDLDGDFVSGDPKQRPQASGILLIRKGLIKNGKE
 SQTLELAAPTFRESLDLAPRLPLAASRPSGAVSDSISSSFOLDLIPERLPEP
 PLRGCVVDNLTLGEPPSSEPGSCLLRWODPLGDSFSLTDQEVATAYRQALRVC
 SKLT"
 ORIGIN
 Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
 Best Local Similarity 81.7%; Pred. No. 0;
 Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
 Qy 82 CTACTGCTGCCGACCACTCACTGCGGCTCCCGCAGCACTGAGTGGTGGCTTAC 141

Db 77 CTAATGATTTCCCTCCGCGCCCTTCTCACTCTCTGCTGCGATGGAAGTCGCTGGTTTCG 136
Qy 142 TCCAGCGCCGAGCCAGGCTCTCTCGGCTTCCGCTCGGCTCGAGAGCGCCGCGCGC 201
Db 137 CGCGCCCTCCGCGCCCACTCTCTCGGCGGAGAGCTA-----GCCGCGCGC 184
Qy 202 TGGCGGAGCGGCTCATCAAGTGGCTTAACTCTGATGTAAGAGAGCGGTTGAAGCGCGC 261
Db 185 TGGCGGAGAGGCTATCAAGTGGCCCAAGCCCTTAATGTAAGAGAGCGGTTGAAGCGCGC 244
Qy 262 ACCATCACAAAACAACCTCGCGGACCGCTACGAGTTCTGAGAGCGCTGGGCAAGGCA 321
Db 245 ACCACCAAGAGCAACAACCTCGCGGACCGCTACGAGTTCTGAGAGCGCTGGGCAAGGCA 304
Qy 322 CTTACGGGAGGTGAAGAGGCAAGAGAGCTCGGGGCGTCTGGTGGCATCAAGTCA 381
Db 305 CTTACGGGAGGTGAAGAGGCGCGGAGAGCTCGGGGCGCTGGTGGCATCAAGTCA 364
Qy 382 TCAAGGAAAGCAAAATGAANAATGAGAGAGTCTGCTCAATACGAGGAGAGTTGAGA 441
Db 365 TCCGGAAGAGCAAAATGAANAATGAGAGAGTCTGATGACATACGAGGAGAGTTGAGA 424
Qy 442 TCAATGCTTCACTCAACCAACCCCAACATCATTTGCCATCATGAGTGTGAGAAATGCA 501
Db 425 TCAATGCTATCACTCAACCAACCCCAACATCATTTGCCATCATGAGTGTGAGAAAGCA 484
Qy 502 GCAAGATTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 485 GCAAGATTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
Qy 562 AGCGGCGACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
Db 545 AGCGGCGACCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604
Qy 622 TGCATCTGCTGCAACAGAGCGGATCGTTCAACGAGATCTCAAGCTGGAAGAAACCTCTC 681
Db 605 TGCATCTGCTGCAACAGAGCGGATCGTTCAACGAGATCTCAAGCTGGAAGAAACCTCTC 664
Qy 682 TAGATGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 741
Db 665 TAGATGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 724
Qy 742 GCAAGTTCTTCCCAAGCGGATCGTTCAACGAGATCTCAAGCTGGAAGAAACCTCTC 801
Db 725 GCAAGTTCTTCCCAAGCGGATCGTTCAACGAGATCTCAAGCTGGAAGAAACCTCTC 784
Qy 802 GGAAGCCCTATGTTGGGCGGCAAGAGTGAAGCTGATGATGATGATGATGATGATGATG 861
Db 785 GGAAGCCCTATGTTGGGCGGCAAGAGTGAAGCTGATGATGATGATGATGATGATGATG 844
Qy 862 TGGTGCATGACACCATGCTCTTTTGAAGGAGATCAATAAACACTGATGAGCAAAATCA 921
Db 845 TGGTGCATGACACCATGCTCTTTTGAAGGAGATCAATAAACACTGATGAGCAAAATCA 904
Qy 922 GTAAAGGCGGCTTACCGTGAAGCGGCAAGCGCTCGGATGCTGAGGCTGATGAGGCTG 981
Db 905 GCAACGGGCGCTTACCGTGAAGCGGCAAGCGCTCGGATGCTGAGGCTGATGAGGCTG 964
Qy 982 TGTATATGATGAACCCACCGGCTGAGGCGACATGAGAGATGAGAGATGATGAGGAG 1041
Db 965 TGTATATGATGAACCCACCGGCTGAGGCGACATGAGAGATGAGAGATGATGAGGAG 1024
Qy 1042 TCAAATGAGGTTTACCAACCGGAGTGGGGAACAGAAAGCCCTGATGAGGAGGAGCAC 1101
Db 1025 TCAAATGAGGTTTACCAACCGGAGTGGGGAACAGAGAGGCTCCGCAATGAGGAGGAGCAC 1084
Qy 1102 CTAGTGTGATCTTTGCGCGGCGCTCATGCGGAGCTGATGATGATGATGATGATGATG 1161
Db 1085 CTGCGAGTGAATCTGCGCGGCGCTCATGCGGAGCTGATGATGATGATGATGATGATG 1144
Qy 1162 TCCGAGGAGATGAGCAAGGTTGTCAGCTTTTCAAGCGAGCAAGTCCCGGAGAGTGA 1221

Db 1145 TCCGAGGAGATGAGGCGCAAGGTGTCAGCTTTCTTCAAGCAGATGCACTGCTGGGGA 1204
Qy 1222 GCACTGATCTGAGGCTGAGAGGCAACATTTCTTAAAGAGTCCGAAAGAGATGACA 1281
Db 1205 GCAACACCCCTGAGGCTGAGAGGCGAGATTCGCTCAAGAAAGTCCCGAAGAGATGACA 1264
Qy 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGGATACCTGTTCTGCGCTGCGCAAGGCA 1341
Db 1265 TGGCCAGTCTCTTCAAGTGAACCGGCTGATGACATGCCATGCGCTGCGCAAGGCA 1324
Qy 1342 GCTTAAAGCTTCCGAAAGCAATTTCAAGAAAGTCTCTTCACTCGTCAAGGAGGATAC 1401
Db 1325 ACTCAAGCTGCGCAAGAGGCAATTTCAAGAAAGTCTGAGCTCTGCGCAAGAGGATAC 1384
Qy 1402 AGAGAGACCTTCAAGAACTCAAGCGGCTGATGATGATGATGATGATGATGATGATG 1461
Db 1385 AGAGAGACCTTCAAGAACTCAAGCGGCTGATGATGATGATGATGATGATGATGATG 1438
Qy 1462 TATCCCTGCTCCCAAGGAAAGGCACTCTTAAAGATGTCAGACAGCTGAATCTGTTACT 1521
Db 1439 ---CCCTGCTCCCAAGGAGGCAATTTCAAGAAAGCCCGCAAGCGGAGTCTGTTACT 1495
Qy 1522 ACTCCTTCCAGAGCCGAGGAGTCTGAGGAACTTTAGACGCGCATGATGTTTGTGA 1581
Db 1496 ACTCCTTCCAGAGCCGAGGAGTCTGAGGAGTCTTGAAGCGAGCGAGTGTGTTGTA 1555
Qy 1582 GTGGGAGACCCGTGAGCAGAAAGTCTTCAAGGCTTCAAGGCTTCTTCCACCGCAAG 1641
Db 1556 GTGGGAGATCCCAAGAGCAGAAAGCTTCCGCAAGGCTTCAAGGCTTCTTCCATGCAAG 1615
Qy 1642 GCATTTCAAAATCTCAATGAGGATGTTCTTCCGCAAGGCTTGAAGGCACTAACCTTAGCA 1701
Db 1616 GCATTTCAAAATCTCAATGAGGATGTTCTTCCGCAAGGCTTGAAGGCACTAACCTTAGCA 1675
Qy 1702 CCTTGGCTCCCTGAGCAACATGAGCTCTTCCCAATCTGAGAGCCGCGCCAGCGGCCCT 1761
Db 1676 CCTTGGCTCCCTGAGCAACATGAGCTCTTCCCAATCTGAGAGCCGCGCCAGCGGCCCT 1735
Qy 1762 CAGGAGGCTGTGATGAGGAGCAGATCTGTTCTTCCGAGTCTTGAACCAATGAGCTTGC 1821
Db 1736 CAGGAGGCTGTGATGAGGAGCAGATCTGTTCTTCCGAGTCTTGAACCAATGAGCTTGC 1795
Qy 1822 CTGAAGCTTTTCCGAAACCCCACTGAGGAGGCTGTGTCTGTGAGCAACCTGAGGAGG 1881
Db 1796 CTGAAGCTTTTCCGAAACCCCACTGAGGAGGCTGTGTGTGTGAGCAACCTGAGGAGG 1855
Qy 1882 TTGAGCAGCTCCCTCGAAG-----GTGGAAGCAGATGATGAGGAGATCT 1929
Db 1856 TTGAGAGAGCCCTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1915
Qy 1930 TGGGAGATGAGCTGTTTCTTCTGACAGATGCAAGAGGATGATGAGGAGGAGGAGGAG 1989
Db 1916 TGGGAGATGAGCTGTTTCTTCTGACAGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1975
Qy 1990 CCTTGAAGATCTGCTCAAGAGTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2049
Db 1976 CACTGAGAGGCTGCTCAAGAGTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2034
Qy 2050 GCTTGAAGAGGTTTGAAGAGGAAACCTGAGTGAATTCCTC 2092
Db 2035 GCTTGAAGAGGTTTGAAGAGGAAACCTGAGTGAATTCCTC 2077

RESULT 12
BD127309
LOCUS BD127309 3395 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127309
VERSION BD127309.1 GI:23222254
KEYWORDS JP 2002017375-A/2740.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 3395)
Oka,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE
JOURNAL
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2740 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2740
PP 22-JAN-2002
PP 07-JUL-2000 JP 2000253172
PI TOSHIO OKA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI YURI KAWAI,AI MAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (118)..(2001).
1.3395
Location/Qualifiers
1.3395
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
QY 82 CTACTGCTGCCGACCCATCTCACCCTGCGGCTCCCGACCATGAGTGGTGGCTTTAC 141
DB 77 CTACTGATTCCTCCGCGCCCTTCTGCTCACTCTGCTGCGCATGAGTGGTGGCTTTAC 136
QY 142 TCCAGGCGCCGACGAGCTCCCTGCGGCTCCCGGCTGCGGCTCGGAGAGCCCGGCGGC 201
DB 137 CGGCGCCCTCCGCGCCCACTCTCTGCGCGGAGAGCTA-----GCCCGCGGC 184
QY 202 TGGCGGAGGAGCTCATGAGTGGCTTAACCTCTGATGAAGAAGAGCGGTGAAGCGGC 261
DB 185 TGGCGGAGGAGCTCATGAGTGGCTTAACCTCTGATGAAGAAGAGCGGTGAAGCGGC 244
QY 262 ACCATCAACAACAACCTGCGGACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGCA 321
DB 245 ACCACCAACAACAACCTGCGGACCGCTACGAGTTCCTGAGAGCGCTGGGCAAGGCA 304
QY 322 CTTACGGGAGGAGTGAAGAAGAGCAGAGAGCTGGGGCGTCTGTGGCCATCAAGTCA 381
DB 305 CTTACGGGAGGAGTGAAGAAGAGCAGAGAGCTGGGGCGTCTGTGGCCATCAAGTCA 364
QY 382 TCAGGAAGAACAACAATCAAGATGAGAGAGTCTGTGCAATACGAGAGGAGATTGGA 441
DB 365 TCCGGAAGGAGCAAAATCAAGATGAGAGAGTCTGTGCAATACGAGAGGAGATTGGA 424
QY 442 TCATGTCTTCACTCAACCAACCCCAATCATTCATTCATCAAGAGTGGTGAATGCA 501
DB 425 TCATGTCTTCACTCAACCAACCCCAATCATTCATTCATCAAGAGTGGTGAATGCA 484
QY 502 GCAAGATTGTGATGTATGATGAGATGAGAGAGTCTGTGATGATGATGATGATGATG 561
DB 485 GCAAGATTGTGATGTATGATGAGATGAGAGAGTCTGTGATGATGATGATGATGATG 544
QY 562 AGCGGCAAGGCTGATGAGAGGAGAGCGGAGCAATTTCTTCCGACAGATGCTGTGCGC 621
DB 545 AGCGGCAAGGCTGATGAGAGGAGAGCGGAGCAATTTCTTCCGACAGATGCTGTGCGC 604
QY 622 TGCACTAATGCTACCAAGAGCGGAGTCTTCAACGAGATCTCAAGCTGGAATAATCTTCT 681

DB 605 TGCACTAATGCTACAGAAAGAGTTCACCGAGATCTCAAGCTGAGAAATCTCTCT 664
QY 682 TAGATGCCAATGGAACATCAAGATTCGATCTTGGCTCTCTCCAACTGTACCAAAAG 741
DB 665 TAGATGCCAATGGAATATCAAGATTCGATCTTGGCTCTCTCCAACTGTACCAAAAG 724
QY 742 GCAAGTTCCTCAGAGCTTCTGTGGAGCCCTCTTACGCTCGCTGAGATGATGCAAG 801
DB 725 GCAAGTTCCTCAGAGCAATTCGTGGAGCCCTCTTATGCTCTCCAGAGATTCATG 784
QY 802 GGAAGCCCTAATGAGGAGCCAGAGGTGAGAGCTGTGCTGGGGCTCTCTGATCAATCC 861
DB 785 GGAAGCCCTAACAAGGAGCCAGAGGTGAGAGCTGTGCTGGGGCTCTCTGATCAATCC 844
QY 862 TGTGTCATGAGCACCATCCCTTTGACCGGAGATCAATAACAATGAGTGAAGCAATCA 921
DB 845 TGTGTCATGAGCACCATCCCTTTGATGAGGATGACATTAAGATCTGATGAAAGATCA 904
QY 922 GTAAAGGAGCTTACCGTGAAGCGCCCAAGCGCTGATGCTGTGCGCTGATCCGCTGCG 981
DB 905 GCAAGGAGGCTTACCGGAGGAGCACTAAACCTCTGATGCTGTGCGCTGATCCGCTGCG 964
QY 982 TGTTAATGATGAACCCCAACCCCTGTGGGCACTGAGAGATGATGATGATGATGATG 1041
DB 965 TGTGATGATGAACCCCAACCCCTGTGGGCACTGAGAGATGATGATGATGATGATG 1024
QY 1042 TCACCTGAGGATTCACCAACCGAGTTCGAGGAGCAAGAGCCCTGATGAGGAGTGGGACC 1101
DB 1025 TCACCTGAGGATTCACCAACCGAGTTCGAGGAGCAAGAGCCCTGATGAGGAGTGGGACC 1084
QY 1102 CTAGTGTGATCTTGGCGGAGCTTCATGAGGAGCTGATGATGATGATGATGATGATG 1161
DB 1085 CTGACAGTGTGATCTTGGCGGAGCTTCATGAGGAGCTGATGATGATGATGATGATG 1144
QY 1162 TCTGAGAGATGAGAGCCCAAGGATGATGATGATGATGATGATGATGATGATGATG 1221
DB 1145 TCTGAGAGATGAGAGCCCAAGGATGATGATGATGATGATGATGATGATGATGATG 1204
QY 1222 GCACTGTACCTGAGGCTGAGAGGAGCAATCTCTTAAGAAGTCCGAAGAAGAGATGCA 1281
DB 1205 GCACTGTACCTGAGGCTGAGAGGAGCAATCTCTTAAGAAGTCCGAAGAAGAGATGCA 1264
QY 1282 TGGCTCAAAATCTGCAAGGATGACCGGCTGAGATGATGATGATGATGATGATGATG 1341
DB 1265 TGGCTCAAAATCTGCAAGGATGACCGGCTGAGATGATGATGATGATGATGATGATG 1324
QY 1342 GCTTAAGTTCGGAAGGATTTCTCAAGAAAAGTCTCTTCACTGTCAGGAGGATGAC 1401
DB 1325 ACCTCAAGTTCGGAAGGATTTCTCAAGAAAAGTCTCTTCACTGTCAGGAGGATGAC 1384
QY 1402 AGAGAGACCTTCAGAGATGACCGGCTGATGATGATGATGATGATGATGATGATG 1461
DB 1385 AGAGAGACCTTCAGAGATGACCGGCTGATGATGATGATGATGATGATGATGATG 1438
QY 1462 TATCCCTGCTCCCAAGGAAAGGATCTTAAAGATCTGCAAGGCTGAAATCTGATGAT 1521
DB 1439 TATCCCTGCTCCCAAGGAAAGGATCTTAAAGATCTGCAAGGCTGAAATCTGATGAT 1495
QY 1522 ACTCTCTCGAAGCCCAAGGATCTTAAAGATCTGCAAGGCTGAAATCTGATGATGAT 1581
DB 1496 ACTCTCTCGAAGCCCAAGGATCTTAAAGATCTGCAAGGCTGAAATCTGATGATGAT 1555
QY 1582 GTGGGAGCCCTGAGAGCAAGATCTCAAGGCTTCAAGGCTCTCTTCAAGGCAAGG 1641
DB 1556 GTGGGAGCCCTGAGAGCAAGATCTCAAGGCTTCAAGGCTCTCTTCAAGGCAAGG 1615
QY 1642 GCATTTCTCAAACTCAATGAGCAAGTTCCTGCAAGGCTTCAAGGCAAGTTCCTTAC 1701
DB 1616 GCATTTCTCAAACTCAATGAGCAAGTTCCTGCAAGGCTTCAAGGCAAGTTCCTTAC 1675
QY 1702 CTTTGGCTCTCTGAGCAAGTGGCTCTCTTCAAGGCTTCAAGGCTTCAAGGCAAG 1761
DB 1676 CTTTGGCTCTCTGAGCAAGTGGCTCTCTTCAAGGCTTCAAGGCTTCAAGGCAAG 1735

QY 1762 CAGGGGCTGAGTGAAGAGACGATCTCTCTCCGAGTCTTTGACCAATTGCACTTGC 1821
| | | | |
DB 1736 CAGGGGCTGAGTGAAGAGACGATCTCTCTCTGATCTTTGACCAAGCTGATCTGC 1795
| | | | |
QY 1822 CTGAACGCTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGGC 1881
| | | | |
DB 1796 CTGAACGCTCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGGC 1855
| | | | |
QY 1882 TTGAGCAGCTCTCCCTCAGAG------GTCTGAGCAGTGTGTGAGGAGATCTCT 1929
| | | | |
DB 1856 TTGAGGAGCCCTCCCTCAGAGGGCCCTGAGAGCTGCTGAGGGCTGTGGGGCAGATCTCT 1915
| | | | |
QY 1930 TGGGGGATGAGTGTCTTTCTCTGACAGCTGCCAAGAGTCACTGACCTTACAGCAAG 1989
| | | | |
DB 1916 TGGGGGACAGCTGTCTTTCTCTGACAGCTGCCAAGAGTCACTGACCTTACAGCAAG 1975
| | | | |
QY 1990 CCCAGAGATCTGCTCAAGAGCTGAGAGAGGAGATGATGCTCCATGATGAGGAG 2049
| | | | |
DB 1976 CACTGAGAGGCTGCTCAAGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2034
| | | | |
QY 2050 GCTCTGAGAGGTTTGAAGAGAACCTGAGTGTGATTCCTCC 2092
| | | | |
DB 2035 GCTCTGAGATGACAGTGTGACACCCGAGGGAGATGCTTC 2077
| | | | |
RESULT 13
AK074830 3395 bp mRNA linear PRI 03-SEP-2002
LOCUS AK074830
DEFINITION Homo sapiens cDNA FLJ90349 f18, clone NT2RP2003179, weakly similar
to PUTATIVE SERINE/THROMBIN-PROTEIN KINASE EMK (EC 2.7.1.1.-).
ACCESSION AK074830
VERSION AK074830.1 GI:22760531
KEYWORDS oligo capping; f18 (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Iisogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahara,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3395)
AUTHORS Iisogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Iisogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo; Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1.3395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2003179"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RP2"
/note="Cloning vector: pMB18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"
118..2004

/note="unnamed protein product"
/codon_start=1
/protein_id="Bac11234.1"
/db_xref="GI:22760532"
/translation="MESLVPARRSGPTPSAALPLAELIKSPKIMKQAVRRH
HKNNLRHYEFLETLGKGYKAKKARESSGRIVAIKSIKDKIKDEODIMHIREIE
IMSLNHPHILAIHEVPENSKIYIWEVYSRDLVDYISEROOLSRERAPRQIV
SAVHYCHONRVHRDLKLENTILDANGNTIKIADPGLSNLTHQGFIDTFCGPIYAS
EIVNGKPYTGEVDSWSLGLVLIIVHNGVAFPGHDKILVKQISNAYRPPKPSDA
CGILRMILMNPTRRATLEVDVASHMMVWNGVYARVGEQEPHGGHSGDSASAPAD
WLRSSPFLLENKAYCSFEKQAPGGSTTPELEOHSLKSKKENDMOSHSIDTA
DPTAHRPKENLTLPGIKKYVASAEVGEQPEPLSP1PASPGQAPLIPKKGILK
KPRORSQYVSSSPSESGELIDAGVYFVSGDKEQKPPASALLHKRIKILKNGE
SOTPLAIAFTTYSLDLAPRPPPLPASPPSASVSDSTLSSFPQDLPERLSP
PLRGCVSDNLTLGLEBPPSEBPSCLRRWQDPLGDSFSLTDCQEVATYROLRVC
SKLT"

Query Match 47.4%; Score 1375.8; DB 9; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
ORIGIN
QY 82 CTAATGCTGCCGACCACTCCACCTGCGCGATCCCGGACCAATGAGTCCGATCCCTAC 141
| | | | |
DB 77 CTAATGATTCCTCCCTCGCGCTTGTCTACCTCTGCTGCGCATGAGTCCGTGTTTCG 136
| | | | |
QY 142 TCAGCGCCGAGCGAAGCTCCTCGCGCTTCGCGCTTCGAGAGCGCCGCGCGC 201
| | | | |
DB 137 CGCGCGCTCCGCGCCCACTCCCTCGCGCCGAGAGCTA-----GCCGCGCGC 184
| | | | |
QY 202 TGGCGGACGGGCTCATTAATCGCTTAACCTCTGATGAAGAGCAGCGGTGAAGCGGC 261
| | | | |
DB 185 TGGCGGAGGGGCTGATCAAGTCCCAAGCCCTTAAGAGAGCAGCGGTGAAGCGGC 244
| | | | |
QY 262 ACCATCAACAACAACCTCGGACCGCTACAGAGTCTCTGAGACGCTGGGCAAGGGCA 321
| | | | |
DB 245 ACCACCAACAACAACCTCGGACCGCTACAGAGTCTCTGAGACGCTGGGCAAGGGCA 304
| | | | |
QY 322 CTAACGGAGAGTGAAGAGCAGAGAGAGCTCGGGGCTCTGTGTGGCCATCAAGTCA 381
| | | | |
DB 305 CTAACGGAGAGTGAAGAGCAGAGAGAGCTCGGGGCTCTGTGTGGCCATCAAGTCA 364
| | | | |
QY 382 TCAGGAAGCAAAATTAAGATGAGAGAGATCTGCTGCAATACGAGGAGATTGAGA 441
| | | | |
DB 365 TCCGGAAGGCAAAATTAAGATGAGAGAGATCTGATGCAATACGAGGAGATTGAGA 424
| | | | |
QY 442 TCATGCTTCACTCAACCAACCCCAATCATCTTGCATCCATGAAGTTTGAATGCA 501
| | | | |
DB 425 TCATGCTATCATCAACCAACCCCAATCATCTTGCATCCATGAAGTTTGAATGCA 484
| | | | |
QY 502 GCAAGATTGATTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 561
| | | | |
DB 485 GCAAGATTGATTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 544
| | | | |
QY 562 AGCGGCAACGCTGAGTGAAGCGGAGCGCAGGCAATTTCTTCCGACAGATGTGTGCC 621
| | | | |
DB 545 AGCGGCAACGCTGAGTGAAGCGGAGCGCAGGCAATTTCTTCCGAGATGTGTGCC 604
| | | | |
QY 622 TGAATCTAGCCCAACGAGAGGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTC 681
| | | | |
DB 605 TGAATCTAGCCCAACGAGAGGATGTTCAACGAGATCTCAAGCTGAGAAACATCTTC 664
| | | | |
QY 682 TAGATGCAATGAGAAACATCAAGATTGCTGATCTTGGCTCTCCAACTGTACACAAAG 741
| | | | |
DB 665 TAGATGCAATGAGAAATATCAAGATTGCTGATCTTGGCTCTCCAACTGTACACAAAG 724
| | | | |
QY 742 GCAAGTTCTTCCAGAGCTTCTGTGGAGCCCTCTTACAGCTTGTGCTGAGATAGTCAAG 801
| | | | |
DB 725 GCAAGTTCTTCCAGAGATTTGTGGAGCCCTCTTACAGCTTGTGCTGAGATAGTCAATG 784
| | | | |
QY 802 GGAAGCCATATGAGGCGCAGAGTGAAGAGCTGCTGAGGCGTTCTCCCTACATCC 861
| | | | |
DB 785 GGAAGCCATACAGGCGCAGAGTGAAGAGCTGCTGAGGCGTTCTCCCTACATCC 844
| | | | |

QY 862 TGGTGCATGACACATGCTCTTTGACGCGGAGATCATTAACAAGCACTGTGAGAAATCA 921
 Db 845 TGGTGCATGACACATGCTCTTTGATGGGATGATCATTAAGATCTTGTGAAACAGATCA 904
 QY 922 GTTAACGGGGCTTACCGTGAAGCCGCCCAAGCGTCCGATGCTGTGGCTGATCCGGTGC 981
 Db 905 GCAACGGGGCTTACCGGGAAGCACTTAACCTCTGTGCTGTGGCTGTGCTGTGCTGCTG 964
 QY 982 TGTAAATGTGAACCCGACCGGTGCGGCGACATGTGAGAGATGTAGCAATGTGAGG 1041
 Db 965 TGTGATGTGAACCCGACCGGTGCGGCGACCTGTGAGAGATGTGAGCACTGTGAGG 1024
 QY 1042 TCAACTGGGGTTACACCAACCGAGTCCGGGAAACAGAAAGCCCTGCTGAGGGTGGGACC 1101
 Db 1025 TCAACTGGGGCTACCGACCCGAGTGGGAGAGAGAGGCTCCGACATGAGGGTGGGACC 1084
 QY 1102 CTAGTGTGATCTTTGCGGGCTCTCATGCGGAGCTGTTAGTGTGCTCTCCGCGCC 1161
 Db 1085 CTGCGAGTGAATCTGCGCGGCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144
 QY 1162 TCCGTGAGATGAGCCGAGGTGTGAGCTTCTTCAAGACAGCATGTGCGGGAGTGA 1221
 Db 1145 TCCGTGAGATGAGCCGAGGTGTGAGCTTCTTCAAGACAGCATGTGCGGGAGTGA 1204
 QY 1222 GCACTGTACCTGGGCTGAGCGGACGACATTTCTTAAGAAAGTCCGAAAGAGATGACA 1281
 Db 1205 GCAACACCCCTGGCGCTGAGCGGACGACATTTCTTAAGAAAGTCCGAAAGAGATGACA 1264
 QY 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGATCCTTCTGCGCTTGGCAAGACA 1341
 Db 1265 TGGCGGAGTCTCTTCAAGTGAACGCGCTGATGACATGCGCCATGCGCTTGGCAAGACA 1324
 QY 1342 GCTTTAAGCTTCCGAAAGGCAATTTCAAGAAAGTCTCTCACTGCTGAGGGAGTAC 1401
 Db 1325 ACCTCAAGCTGCAAAAGGCAATTTCAAGAAAGTCTCTCACTGCTGAGGGAGTAC 1384
 QY 1402 AGGAGACCTCTGAGAACTGACAGCGGCTGCTGATCTCAAGGAGCAGCTGCTGCTG 1461
 Db 1385 AGGAGACCTCTGAGAACTGACAGCGGCTGCTGATCTCAAGGAGCAGCTGCTGCTG 1438
 QY 1462 TATCCCTGCTCCAGAAAGGCAATTTCAAGAAAGTCTCTCAAGGAGCAGCTGCTGCTG 1521
 Db 1439 ---CCCTGCTCCAGAAAGGCAATTTCAAGAAAGCCTGAGCGGAGTCTGCTGCTGCTG 1495
 QY 1522 ACTGCTCTCAAGAGCCGAGAGTCTGCGGAACTTTAGAGCCAGTGTGTTGTGA 1581
 Db 1496 ACTGCTCTCAAGAGCCGAGAGTCTGCGGAACTTTAGAGCCAGTGTGTTGTGA 1555
 QY 1582 GTGGGAGCCCGTGGAGCAAGATCTCAAGGCTTCAAGGCTCTGCTCAAGGAGG 1641
 Db 1556 GTGGGAGTCCAGAGCAAGATCTTCAAGGCTTCAAGGCTCTGCTCAAGGAGG 1615
 QY 1642 GCATTTCAAACTCAATGCAAGTCTTCCGCAAGCCTTGAAGGCACTAACCTTGA 1701
 Db 1616 GCATTTCAAACTCAATGCAAGTCTTCCGCAAGCCTTGAAGGCACTAACCTTGA 1675
 QY 1702 CTTTGGCTCTGAGCAAACTGCGCTCTCCATCTGAGCCGCGCCAGCCGCT 1761
 Db 1676 CTTTGGCTCTGAGCAAACTGCGCTCTCCATCTGAGCCGCGCGCCAGCCGCT 1735
 QY 1762 CAGGGGCTGAGGAGAGAGAGATCTGCTGCTCCGAGTCTTGAACAATGTGAGCT 1821
 Db 1736 CAGGGGCTGAGGAGAGAGAGATCTGCTGCTCCGAGTCTTGAACAATGTGAGCT 1795
 QY 1822 CTGAACGTCTTCCGAAACCCCACTGAAGGGCTGTGTGTGTGAGCAAACTGAGGGG 1881
 Db 1796 CTGAACGTCTTCCGAAACCCCACTGAAGGGCTGTGTGTGTGAGCAAACTGAGGGG 1855
 QY 1882 TTGAGAGGCTCTCTGAGAG---GTCTGAAGCATGTGTGAGAGAACTCT 1929
 Db 1856 TTGAGAGGCTCTCTGAGAGGCTCTGAGAGGCTCTGAGAGGCTCTGAGAGATCTCT 1915

QY 1930 TGGGGATACCTGCTTTTCTCTGACAGACTGCGCAAGAGTGTGATGCACTTACAGACAG 1989
 Db 1916 TGGGGATACCTGCTTTTCTCTGACAGACTGCGCAAGAGTGTGATGCACTTACAGACAG 1975
 QY 1990 CCTAGAAATCTCTCAAAAGCTGAGAGGAGGAGAGTGTGCTGATGATGAGGGTAG 2049
 Db 1976 CACTGAGGGTCTCTCAAAAGCTGAGAGGAGGAGAGTGTGATGATGAGGGTAG 2034
 QY 2050 GCTCTGAGAGGTTTGGAGAGAACCTGCGTGGATTCCTCC 2092
 Db 2035 GCTCTGAGAGTGTGAGTGTGACCCGAGGAGAGATGCTTC 2077
 RESULT 14
 BC017306
 LOCUS
 DEFINITION
 Homo sapiens likely ortholog of rat SNR1/AMP-activated protein kinase, mRNA (CDNA clone MGC:29812 IMAGE:5088037), complete cde.
 BC017306
 VERSION
 ACCESSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K., Hopkins, R.F., Jordan, J.A., Moore, T., Max, S.L., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Gilmour, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, J.M., Skalski, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3402)
 Strausberg, R.
 Direct Submission
 Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:16878205.
 Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: Lou Staudt
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 Sequencing Center (NISC),
 Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Ahlter, N., Ayala, K., Beckstrom-Stenborg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hashiguchi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaardi, R., Maduro, Q.L., Mastello, C., Maekari, B., Maestrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,

Teurjeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L., Young, A., Zhang, L.-H. and Green, R. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 39 Row: 0 Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 13565921.

FEATURES

Source

1..3402

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:29812 IMAGE:5088037"

/tissue_type="lymph, lymphoma"

/clone_id="NH MGC_99"

/lab_host="DH10B"

/note="Vector: pOTB7"

1..3402

/gene="SNARK"

/note="synonym: DKFZP434J037"

/db_xref="LocustID:81788"

/db_xref="MIM:608131"

100..1986

/gene="SNARK"

/codon_start=1

/product="hypothetical protein DKFZP434J037"

/protein_id="AAH17306.1"

/db_xref="GI:16878206"

/db_xref="LocustID:81788"

/db_xref="MIM:608131"

/translation="MESIVAPRPSGTPPSAAELARPIAEGLIKSPKLMKKQAVKRH
HKNNRRYFLETLGKTYGKAPKABESSRLVAISIRKDKIDQDMHIREIE
IMSLNPHI IAHVEPENSCKIVMEVASRGDLYDIERQDSREARHFRQIV
SAHYGQONVHNDLKENILLDANGNIKIDPGLNHYOGFTQFCGSPLYASP
EIVNGKYTPVNDMSLGVLLYLVAHTGTFDGHDKILVKQISNCAVREPPKPSA
CGILRLIMVNPTRATLEDOVASHMVMWGYATRVGROBA PHEGGHSGDSASASMD
WLRSSPILNENKAYCSPPKQAH RGGSTTPGLERPHSLKSKKNDMAQSLSDPA
DTPAHRGKSNLKPKEILKRTVSASABGVQDDPPLSPITPASFGQAPLLPKGIILK
KPRQREGYVSSPEPSGELLDAGDVFSGDPEKQRPQAAGLILHRKQILIKNGF
SQIALBIAPPTTSGSLDELAPRPLAPASPSGAVSDSILSSSPQDLPERLPP
PLRGVSVNDLITGLEBPSBSPGSLARMHQDPLGDSCFSLDQEVATATYRQALRYC
SKLT"

ORIGIN

Query Match 47.4%; Score 1375.8; DB 9; Length 3402;

Best Local Similarity 81.7%; Pred. No. 0;

Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;

QY 82 CTAAGTGTGCGGACCACTGCACTGCGGCTCCCCGACCATGAGTGGTGGCTTAC 141
DB 59 CTAAGTGTGCGGACCACTGCACTGCGGCTCCCCGACCATGAGTGGTGGCTTAC 118
QY 142 TCCAGCGCCGAGCGGACCTCCCTCGGCTCCGCGCTGGCCTCGAGAGCGCCGCGCGC 201
DB 119 CGGCGCGCTCCGCGCGGACCTCCCTCGGCTCCGCGCTGGCCTCGAGAGCGCCGCGCGC 166
QY 202 TGGCGGAGCGGCTCATCAAGTGGCTTAACCTCTGATGAGAGAGCAGCGGTGAGAGCGC 261
DB 167 TGGCGGAGCGGCTCATCAAGTGGCTTAACCTCTGATGAGAGAGCAGCGGTGAGAGCGC 226
QY 262 ACCATCAAAACACCACTGCGGACCGGCTGCGAGTCTCGGAGACGCTGGGCAAGGCA 321
DB 227 ACCATCAAAACACCACTGCGGACCGGCTGCGAGTCTCGGAGACGCTGGGCAAGGCA 286
QY 322 CTTACGGGAGGAGTGAAGAGCAGAGAGCTCGGGGCGTCTGAGTGCATCAAGTCCA 381
DB 287 CTTACGGGAGGAGTGAAGAGCAGAGAGCTCGGGGCGTCTGAGTGCATCAAGTCCA 346
QY 382 TCAAGAAAGCAAAATCAAGATGAGAGGATCTGTGCATATCGAGGAGGATTTAGA 441
DB 347 TCGGAGAGGCAAAATCAAGATGAGAGGATCTGTGCATATCGAGGAGGATTTAGA 406

QY 442 TCATGCTTCACTGAACCAACCCCAATCATTTGCCATTCACATGAAGTGTGAGATATGA 501
DB 407 TCATGCTTCACTGAACCAACCCCAATCATTTGCCATTCACATGAAGTGTGAGATATGA 466
QY 502 GCAAGATTTGATTTGTCATGAGATATGACCGGAGCGAGTCTGTATGATTAATCATCATG 561
DB 467 GCAAGATTTGATTTGTCATGAGATATGACCGGAGCGAGTCTGTATGATTAATCATCATG 526
QY 562 AGCGGCAAGGCTGAGTGAAGCGGAGCGGACGATTTCTTCCAGAGATGTGTGCGCC 621
DB 527 AGCGGCAAGGCTGAGTGAAGCGGAGCGGACGATTTCTTCCAGAGATGTGTGCGCC 586
QY 622 TGACCTATGCGACCAAGAGGAGTGGTTCACGAGATCTCAAGCTGAGAAACATCTCTTC 681
DB 587 TGACCTATGCGACCAAGAGGAGTGGTTCACGAGATCTCAAGCTGAGAAACATCTCTTC 646
QY 682 TAGATGCCAATGGAACATCAAGATTTGCTGATCTTGGCTCTCCACCTGTACCAAG 741
DB 647 TAGATGCCAATGGAACATCAAGATTTGCTGATCTTGGCTCTCCACCTGTACCAAG 706
QY 742 GCAAGTTCCTCAGACGTTCTGTGGAGCCCTCTTACGCTCTGAGATGATCAACG 801
DB 707 GCAAGTTCCTCAGACGTTCTGTGGAGCCCTCTTACGCTCTGAGATGATCAACG 766
QY 802 GGAAGCCCTATGAGGCGCCAGAGTGAACAGCTGTCTGTGGGCTTCTCTGACATCC 861
DB 767 GGAAGCCCTATGAGGCGCCAGAGTGAACAGCTGTCTGTGGGCTTCTCTGACATCC 826
QY 862 TGGTGCAATGACCAATGCTGAGCGGAGAGTCAATAAACACTGAGTGAAGCAATCA 921
DB 827 TGGTGCAATGACCAATGCTGAGCGGAGAGTCAATAAACACTGAGTGAAGCAATCA 886
QY 922 GTAAAGGAGCTTACCGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 981
DB 887 GCAAGCGGAGCTTACCGTGAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 946
QY 982 TGTGTAATGAGTGAACCCCAACCGGCGGAGCACTGAGAGATGAGCAAGTCAATTTGG 1041
DB 947 TGTGTAATGAGTGAACCCCAACCGGCGGAGCACTGAGAGATGAGCAAGTCAATTTGG 1006
QY 1042 TCAACTGAGGCTTACCAACCGGAGTGAAGAGTGAAGAGCGGAGTGAAGAGTGAAG 1101
DB 1007 TCAACTGAGGCTTACCAACCGGAGTGAAGAGTGAAGAGCGGAGTGAAGAGTGAAG 1066
QY 1102 CTAAGTGTGCACTTTGCGGAGCGGAGCTTCAATGAGAGTGGTGGCTCTCGCGCC 1161
DB 1067 CTGAGAGTGTGCACTTTGCGGAGCGGAGCTTCAATGAGAGTGGTGGCTCTCGCGCC 1126
QY 1162 TCTGAGAGATGAGAGCGGAGTGTGAGCTTTTCAAGAGCAAGTGCAGGAGTGA 1221
DB 1127 TCTGAGAGATGAGAGCGGAGTGTGAGCTTTTCAAGAGCAAGTGCAGGAGTGA 1186
QY 1222 GCACTGATCTGAGGCTGAGAGCGGAGCAATCTCTTAAGAGTCCCGAAGAGAGATGACA 1281
DB 1187 GCACTGATCTGAGGCTGAGAGCGGAGCAATCTCTTAAGAGTCCCGAAGAGAGATGACA 1246
QY 1282 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTCTTCTCGGCTTGGAGAGCA 1341
DB 1247 TGGCTCAAAATCTGCAAGTGAACCGGCTGAGATCTCTTCTCGGCTTGGAGAGCA 1306
QY 1342 GCTTAAAGCTTCCGAAAGGATTTCTCAAGAAAAGTCTTACTGCTCAAGGAGGATAC 1401
DB 1307 GCTTAAAGCTTCCGAAAGGATTTCTCAAGAAAAGTCTTACTGCTCAAGGAGGATAC 1366
QY 1402 AGAGAGACCTCTCAGAACTGAGACCGGTGCTGATCTCAGAGGAGCGGCTGCTG 1461
DB 1367 AGAGAGACCTCTCAGAACTGAGACCGGTGCTGATCTCAGAGGAGCGGCTGCTG 1420
QY 1462 TATCCCTGCTCCCAAGAGAGGATCTTAAAGAGTCTGACAGCGGATTTGTTACT 1521
DB 1421 TATCCCTGCTCCCAAGAGAGGATCTTAAAGAGTCTGACAGCGGATTTGTTACT 1477
QY 1522 ACTCTCTTCCAGAGCGGAGGATCTGGGAACTTTAGAGCGCAAGTATGTGTTGTA 1581

Db 1478 ACTCTCTCCGAGAGGAGTAACTGAGGAGCTCTTGAGCAGAGGAGCCTGTTTGA 1537
Qy 1582 GTGGGAGACCCGAGAGAGAGTCTCAGAGGCTTCAAGGCTCTCTCCACCGCAAG 1641
Db 1538 GTGGGAGATCCAGAGAGAGAGGCTCTCAGAGGCTTCAAGGCTCTCTCCATGCAAG 1597
Qy 1642 GCATTCTCAACTCAATGAGCAAGTCTCCGAGAGGCTTGAAGGCACTACCTTGA 1701
Db 1598 GCATCTCAAACTCAATGAGCAAGTCTCCGAGAGGCTTGAAGGCTTGAAGGCACTAC 1657
Qy 1702 CTTTGGCTCTCTGAGCAAGTCTCTCTCAATCTGAGAGGCTTGAAGGCTTGAAGG 1761
Db 1658 CTTTGGCTCTCTGAGCAAGTCTCTCTCAATCTGAGAGGCTTGAAGGCTTGAAGG 1717
Qy 1762 CAGGGGCTGAGAGAGAGAGAGGCTCTGAGAGGCTTGAAGGCTTGAAGGCTTGAAG 1821
Db 1718 CAGGGGCTGAGAGAGAGAGAGGCTCTGAGAGGCTTGAAGGCTTGAAGGCTTGAAG 1777
Qy 1822 CTGAACGCTCTCTCCGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1881
Db 1778 CTGAACGCTCTCTCCGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1837
Qy 1882 TTGAGCAGGCTCTCTCAGAGAG-----GCTTGAAGGCTTGAAGGCTTGAAG 1929
Db 1838 TTGAGCAGGCTCTCTCAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1897
Qy 1930 TGGGGGATGAGGCTTCTTCTGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 1989
Db 1898 TGGGGGATGAGGCTTCTTCTGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 1957
Qy 1990 CCTGAGGATCTGCTCAAGAGCTGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 2049
Db 1958 CACTGAGGCTGCTCAAGAGCTGAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 2016
Qy 2050 GCTTGAAGGCTTGAAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 2092
Db 2017 GCTTGAAGGCTTGAAGAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 2059

RESULT 15
AX086949 3443 bp DNA linear PAT 09-MAR-2001
LOCUS AX086949 Sequence 901 from Patent WO0112659.
DEFINITION AX086949
ACCESSION AX086949
VERSION AX086949.1 GI:13276064
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. Wiemann, S.
AUTHORS Human dna sequences
JOURNAL Patent: WO 0112659-A 901 22-FEB-2001;
German Human Genome Project (DB)
FEATURES
Source 1..3443
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 47.4%; Score 1374.2; DB 6; Length 3443;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy 82 CTACTGCTGCGGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 141
Db 87 CTACTGCTGCGGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 146
Qy 142 TCCAGCGCCGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 201

Db 147 CGCGGCGCTCCGCGCCCACTCTCGGCGGAGAGCTA-----GCCGCGCGCC 194
Qy 202 TGGGAGAGGAGGCTGATCAAGTCCGCTTAACCTGATGAGAGAGAGGCTGAGAGGCGC 261
Db 195 TGGGAGAGGAGGCTGATCAAGTCCGCGCCCAAGCCCTTAATGAGAGAGAGGCTGAGAGGCGC 254
Qy 262 ACCATCAAAACAAACCTTGGGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 321
Db 255 ACCATCAAAACAAACCTTGGGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 314
Qy 322 CTTAAGGAGAGGAGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 381
Db 315 CTTAAGGAGAGGAGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 374
Qy 382 TCAGAGAGAGAGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 441
Db 375 TCAGAGAGAGAGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 434
Qy 442 TCATGCTTCACTCAACCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 501
Db 435 TCATGCTTCACTCAACCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 494
Qy 502 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 495 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 554
Qy 562 AGCGGAGAGGCTGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 621
Db 555 AGCGGAGAGGCTGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 614
Qy 622 TGCACTATGCTCAACCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 681
Db 615 TGCACTATGCTCAACCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 674
Qy 682 TAGATGCAATGGAACATCAAGTCTGATGATGATGATGATGATGATGATGATGATG 741
Db 675 TAGATGCAATGGAACATCAAGTCTGATGATGATGATGATGATGATGATGATGATG 734
Qy 742 GCAAGTCTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 801
Db 735 GCAAGTCTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 794
Qy 802 GGAAGCCCTTATGAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 861
Db 795 GGAAGCCCTTATGAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 854
Qy 862 TGGTCAATGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 921
Db 855 TGGTCAATGAGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 914
Qy 922 GTAAAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 981
Db 915 GTAAAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 974
Qy 982 TGTAAAGGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1041
Db 975 TGTAAAGGAGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1034
Qy 1042 TCAACTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1101
Db 1035 TCAACTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1094
Qy 1102 CTAGTGTGACTTGGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1161
Db 1095 CTAGTGTGACTTGGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1154
Qy 1162 TCTGAGAGAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1221
Db 1155 TCTGAGAGAGGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1214
Qy 1222 GCACTGTACTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1281
Db 1215 GCACTGTACTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 1274


```
Db      20 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTOL-SSSSLOQLFREVRIMKVLNHPN 78
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      79 IYKLFVITTEKTLVYMEYASGGEVFDLVVHGRNKEAKRPFQIYSAVOYCHQKFI 138
Qy      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      139 VHRDLKAENILLDADWNIKIADFGFSENEFTFGNKLDTPFGSPPYAAPBELFQGGKYDGPEV 198
Db      180 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 238
      199 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 258
Qy      239 RATLEDVASHMWVN 252
      259 RGTLEQIMKDRMWN 272
```

RESULT 11

```
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6432156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 722;
Best Local Similarity 51.2%; Pred. No. 5.5e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

Qy      1 YERLELTGKGTGKVKKARE-SSGRVIAKSIKDKIKDQDLHRIREIEMSSLNHPN 59
      53 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTOL-SSSSLOQLFREVRIMKVLNHPN 111
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      79 IYKLFVITTEKTLVYMEYASGGEVFDLVVHGRNKEAKRPFQIYSAVOYCHQKFI 171
Db      112 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      172 VHRDLKAENILLDADWNIKIADFGFSENEFTFGNKLDTPFGSPPYAAPBELFQGGKYDGPEV 231
Qy      180 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 238
      232 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 291
Db      239 RATLEDVASHMWVN 252
      292 RGTLEQIMKDRMWN 305
```

RESULT 12

```
US-10-274-194-4
; Sequence 4, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```

```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-274-194-4
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 722;
Best Local Similarity 51.2%; Pred. No. 5.5e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;
```

```
Qy      1 YERLELTGKGTGKVKKARE-SSGRVIAKSIKDKIKDQDLHRIREIEMSSLNHPN 59
      53 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTOL-SSSSLOQLFREVRIMKVLNHPN 111
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
      79 IYKLFVITTEKTLVYMEYASGGEVFDLVVHGRNKEAKRPFQIYSAVOYCHQKFI 171
Db      112 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      120 VHRDLKLENILLDANGNIKIADFGLSNLVHKGKFLQTFGSPLYASPEIYNGKPYGPEV 179
      172 VHRDLKAENILLDADWNIKIADFGFSENEFTFGNKLDTPFGSPPYAAPBELFQGGKYDGPEV 231
Qy      180 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 238
      232 DWSLGLVLLIYLVHGMPPDGDHKTLLVQISNGAVREPPKPSDAC-GLIRMLMNPTR 291
Db      239 RATLEDVASHMWVN 252
      292 RGTLEQIMKDRMWN 305
```

RESULT 13

```
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36
```

```
Query Match      48.2%; Score 647.5; DB 4; Length 745;
Best Local Similarity 51.2%; Pred. No. 5.8e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;
```

```
Qy      1 YERLELTGKGTGKVKKARE-SSGRVIAKSIKDKIKDQDLHRIREIEMSSLNHPN 59
      20 YRLKLTIGKGNFAKVLAHRIILTGKEVAVKIIDKTOL-SSSSLOQLFREVRIMKVLNHPN 78
      60 IIAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGI 119
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comogen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 14:16:15 ; Search time 197.764 Seconds
(without alignments)
1494.237 Million cell updates/sec

Title: US-09-980-464-11_COPY_57_309

Perfect score: 1342
Sequence: 1 YEFLETLGKGYGVKVKARE.....VNPTRRATLEDVASHMVMVM 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep:*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep:*
33: /cgn2_6/ptodata/1/paa/US107_COMB.pep:*
34: /cgn2_6/ptodata/1/paa/US108_COMB.pep:*
35: /cgn2_6/ptodata/1/paa/US109_COMB.pep:*
36: /cgn2_6/ptodata/1/paa/US110_COMB.pep:*
37: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	1 PCT-US00-14696-11	Sequence 11, Appl
2	1342	100.0	631	25 US-09-980-464-11	Sequence 11, Appl
3	1342	100.0	631	32 US-10-679-063-17712	Sequence 17712, A
4	1342	100.0	631	32 US-10-679-063-19563	Sequence 19563, A
5	1325	98.7	251	29 US-10-343-514-103	Sequence 103, Appl
6	1313	97.8	630	29 US-10-343-514-41	Sequence 41, Appl
7	1313	97.8	630	32 US-10-679-063-20074	Sequence 20074, A
8	1293	96.3	251	29 US-10-343-514-50	Sequence 50, Appl
9	1293	96.3	628	1 PCT-US02-05095A-1163	Sequence 1163, Ap
10	1293	96.3	628	20 US-09-611-526-2919	Sequence 2919, Ap
11	1293	96.3	628	20 US-09-641-371-902	Sequence 902, Appl
12	1293	96.3	628	25 US-09-963-159-2	Sequence 2, Appl1
13	1293	96.3	628	27 US-10-170-205E-24082	Sequence 24082, A
14	1293	96.3	628	29 US-10-370-715B-640	Sequence 640, Appl
15	1293	96.3	628	30 US-10-423-543-44	Sequence 44, Appl
16	1293	96.3	628	32 US-10-679-063-8699	Sequence 8699, Ap
17	1293	96.3	628	35 US-10-932-349-1502	Sequence 1502, Appl
18	1293	96.3	628	37 US-60-452-680-12460	Sequence 12460, A
19	1293	96.3	628	37 US-60-500-337-1502	Sequence 1502, Ap
20	1293	96.3	672	1 PCT-US03-21730-77	Sequence 77, Appl
21	1293	96.3	672	32 US-10-618-941-77	Sequence 77, Appl
22	1286	95.8	594	29 US-10-311-034-12	Sequence 12, Appl
23	1286	95.8	594	37 US-60-216-605-2	Sequence 2, Appl1
24	1276	95.1	251	29 US-10-343-514-101	Sequence 101, Appl
25	1273	94.9	640	29 US-10-322-281-23	Sequence 23, Appl
26	1268.5	94.5	629	25 US-09-979-167-159	Sequence 159, Appl
27	1230	91.7	616	29 US-10-322-281-26	Sequence 26, Appl
28	1128.5	84.1	661	1 PCT-US03-02588-82	Sequence 82, Appl
29	1128.5	84.1	661	1 PCT-US03-38193-2553	Sequence 2553, Ap
30	1128.5	84.1	661	22 US-09-780-949-2	Sequence 2, Appl1
31	1128.5	84.1	661	22 US-09-780-949-6	Sequence 6, Appl1
32	1128.5	84.1	661	22 US-09-791-531-19112	Sequence 149112, A
33	1128.5	84.1	661	27 US-10-170-205E-24527	Sequence 24527, A
34	1128.5	84.1	661	29 US-10-354-358-82	Sequence 82, Appl
35	1128.5	84.1	661	31 US-10-505-928-690	Sequence 690, Appl
36	1128.5	84.1	661	32 US-10-631-067-583	Sequence 583, Appl
37	1128.5	84.1	661	32 US-10-679-063-26806	Sequence 26806, A
38	1128.5	84.1	661	33 US-10-723-860-2553	Sequence 2553, Ap
39	1128.5	84.1	661	35 US-10-932-349-1564	Sequence 1564, Ap
40	1128.5	84.1	661	35 US-10-932-349-1565	Sequence 1565, Ap
41	1128.5	84.1	661	37 US-60-389-395-2	Sequence 2, Appl1
42	1128.5	84.1	661	37 US-60-500-337-1564	Sequence 1564, Ap
43	1128.5	84.1	661	37 US-60-500-337-1565	Sequence 1565, Ap
44	1120.5	83.5	658	1 PCT-US04-14421-150	Sequence 150, Appl
45	1120.5	83.5	658	27 US-10-144-779-497	Sequence 497, Appl

ALIGNMENTS

RESULT 1
PCT-US00-14696-11
Sequence 11, Application PC/TUS0014696
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Virca, George Duke
APPLICANT: Martin, Unja
TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
FILE REFERENCE: 2923-WO
CURRENT APPLICATION NUMBER: PCT/US00/14696
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 631
TYPE: PRT
ORGANISM: Mus musculus
PCT-US00-14696-11

```
Query Match      100.0%; Score 1342; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
QY 61 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 117 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
DB 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
QY 181 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 237 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309

RESULT 2
US-09-980-464-11
; Sequence 11, Application US/09980464
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-US
; CURRENT APPLICATION NUMBER: US/09/980,464
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-980-464-11

Query Match      100.0%; Score 1342; DB 25; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
QY 61 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 117 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
DB 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
QY 181 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 237 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309

RESULT 3
```

```
US-10-679-063-17712
; Sequence 17712, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 17712
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-679-063-17712

Query Match      100.0%; Score 1342; DB 32; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
QY 61 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 117 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
DB 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
QY 181 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 237 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309

RESULT 4
US-10-679-063-19563
; Sequence 19563, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 19563
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-679-063-19563

Query Match      100.0%; Score 1342; DB 32; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.1e-110;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 60
DB 57 YEFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEODLHIREIEMSSLNHPH 116
QY 61 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
DB 117 IAIHEVFENSSKIIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176
QY 121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 180
DB 177 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIYNGKPYGPEVD 236
QY 181 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 240
DB 237 SMSLGLVLYILVHGTMPPFDGQDHKTIVKOISNGAYREPPKPSDAGLIRMLMVNPTTRA 296
QY 241 TLEDVASHMWVNM 253
DB 297 TLEDVASHMWVNM 309
```

Db 177 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 236
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 240
Db 237 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 236
Qy 241 TLEDVASHMWVNM 253
Db 297 TLEDVASHMWVNM 309

RESULT 5

US-10-343-514-103
; Sequence 103, Application US/10343514
; GENERAL INFORMATION:

APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEFEVRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT
CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 103
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-10-343-514-103

Query Match 98.7%; Score 1325; DB 29; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.1e-109;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 60
Db 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 60
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
Db 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
Qy 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Db 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 240
Db 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 240
Qy 241 TLEDVASHMWVNM 251
Db 241 TLEDVASHMWVNM 251

RESULT 6

US-10-343-514-41
; Sequence 41, Application US/10343514
; GENERAL INFORMATION:

APPLICANT: DRUCKER, Daniel J.
APPLICANT: ROSEN, Cheryl F.
APPLICANT: LEFEVRE, Diana L.
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
FILE REFERENCE: DPA-DRUC2/PCT

CURRENT APPLICATION NUMBER: US/10/343,514
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: PCT/CA01/01109
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/222,650
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/274,613
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: CA 2,340,780
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.0
SEQ ID NO 41
LENGTH: 630
TYPE: PRT
ORGANISM: RAT
US-10-343-514-41

Query Match 97.8%; Score 1313; DB 29; Length 630;
Best Local Similarity 98.0%; Pred. No. 8.2e-108;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 60
Db 57 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 116
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176
Qy 121 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 180
Db 177 HRDLKLENTLLDANGNIKIADFGLSNLYHKGKFLQTFCCSPLYASPEIVNGKPYGPEVD 236
Qy 181 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 240
Db 237 SMSLGLVLYLVHGTMPFGQDHKTLYVKQISNGAYREPPSPDAGCGLIRLWLVNPTRR 296
Qy 241 TLEDVASHMWVNM 253
Db 297 TLEDVASHMWVNM 309

RESULT 7

US-10-679-063-20074
; Sequence 20074, Application US/10679063
; GENERAL INFORMATION:

APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52054)B
CURRENT APPLICATION NUMBER: US/10/679,063
CURRENT FILING DATE: 2003-10-02
PRIOR APPLICATION NUMBER: 60/415,758
PRIOR FILING DATE: 2002-10-02
NUMBER OF SEQ ID NOS: 27373
SEQ ID NO 20074
LENGTH: 630
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-679-063-20074

Query Match 97.8%; Score 1313; DB 32; Length 630;
Best Local Similarity 98.0%; Pred. No. 8.2e-108;
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 60
Db 57 YEFLETLGKGTGYKVKKARESSGRLVAIKSIRKDKIKDQDLHIRREIIMSSLNHPH 116
Qy 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 120
Db 117 IAIHEVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIV 176

QY	121	HRDLKLENIILDDANGNIK1ADFGSLNYLHKAGKLTQFCCGSP1LAASEIYVNGKRYGVEVD	180
Db	177	HRDLKLENIILDDASGNIRK1ADFGSLNYLHKAGKLTQFCCGSP1LAASEIYVNGKRYGVEVD	236
QY	181	SMSLGVLIIYILVHGMPFPGDGHKTLVKQISNGAYRBPKEPSNACGLIRLMLVMNPFRR	240
Db	237	SMSLGVLIIYILVHGMPFPGDGHKTLVKQISSGAYRBPKEPSNACGLIRLMLVMNPFRR	296
QY	241	TLEDVASHMVMNM	253
Db	297	TLEDVASHMVMNM	309

RESULT 8
US-10-343-514-50

1 APPLICANT: DRUCKER, Daniel J.
 2 APPLICANT: ROSEN, Cheryl F.
 3 APPLICANT: LEEBEHR, Diana L.
 4 TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
 5 FILE REFERENCE: DPA-DRUC2/PCT
 6 CURRENT APPLICATION NUMBER: US/10/343,514
 7 CURRENT FILING DATE: 2003-01-31
 8 PRIOR APPLICATION NUMBER: PCT/CA01/01109
 9 PRIOR FILING DATE: 2001-08-02
 10 PRIOR APPLICATION NUMBER: US 60/222,650
 11 PRIOR FILING DATE: 2000-08-03
 12 PRIOR APPLICATION NUMBER: US 60/274,613
 13 PRIOR FILING DATE: 2001-03-12
 14 PRIOR APPLICATION NUMBER: CA 2,340,780
 15 PRIOR FILING DATE: 2001-03-28
 16 NUMBER OF SEQ ID NOS: 109
 17 SOFTWARE: PatentIn version 3.0
 18 SEQ ID NO 50

Query Match	96.6%	Score 1296	DB 29	Length 251
Best Local Similarity	98.0%	Pred. No. 8.2e-107		
Matches 246	Conservative 2	Mismatches 3	Indels 0	Gaps 0

QY	1	YEFLETLKGYTGKAYKKAESSGRVAIAKSIKRDOKI	KDEODLHIREIETIMSLNPHI	60
Db	1	YEFLETLKGYTGKAYKKAESSGRVAIAKSIKRDOKI	KDEODLHIREIETIMSLNPHI	60
QY	61	IAIHVEFENNISKIVIMEYASRGDLYDI	SERPRLESDARHFEKROI	VSALHYCHONGIV 120
Db	61	IAIHVEFENNISKIVIMEYASRGDLYDI	SERPRLESDARHFEKROI	VSALHYCHONGIV 120
QY	121	HRDLKLENIILLDANONIKIADRGLSNLVYHKGRFL	OTFCOSPITYASPEIVNGKRYVPEVD	180
Db	121	HRDLKLENIILLDANONIKIADRGLSNLVYHKGRFL	OTFCOSPITYASPEIVNGKRYVPEVD	180
QY	181	SMSLGVLVLIYLVHGTMPEQGDHKTLLVKROI	ISNGAYEPEPKPSDACGLIRWLLVMNPTTRA	240
Db	181	SMSLGVLVLIYLVHGTMPEQGDHKTLLVKROI	ISNGAYEPEPKPSDACGLIRWLLVMNPTTRA	240
QY	241	TLEDVASHMNV	251	
Db	241	TLEDVASHMNV	251	

```

RESULT 9
PCT-US02-05095A-1163
; Sequence 1163, Application PC/TUS0205095A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-075 (803)
; CURRENT APPLICATION NUMBER: PCT/US02/05095A

```

```

1  CURRENT FILING DATE: 2002-03-05
2  PRIOR APPLICATION NUMBER: 09/7799,455
3  PRIOR FILING DATE: 2001-03-05
4  NUMBER OF SEQ ID NOS: 1896
5  SEQ ID NO 1163
6  LENGTH: 628
7  TYPE: prt
8  ORGANISM: Homo Sapiens
9  PCT-OS02-05095A-1163

```

Query Match	96.3%;	Score 1293;	DB 1;	length 628;
Best Local Similarity	95.7%;	Pred. No. 5e-106;		
Matches 242;	Conservative 6;	Mismatches 5;	Indels 0;	Gaps 0;

Qy	1	YEELEFLLGKGTGKVKKARSSGRLVAKSIRKOKIXEODLHIREIIMSLSNPHI	60
Dh	53	YEELEFLLGKGTGKVKKARSSGRLVAKSIRKOKIXEODLHIREIIMSLSNPHI	112
Qy	61	IAHHEFENSSKIVYMEYASRGDLYDISRPPLSERDARHFRQIVSALHYCHONGIV	120
Dh	113	IAHHEFENSSKIVYMEYASRGDLYDISRQOLSEREAHFRQIVSAHYCHONRVV	172
Qy	121	HRDLKLENTLLDANGNIKIADPGLSNLHYHKKFLOTFCGSPLYASPEIVNGKPYGEVD	180
Dh	173	HRDLKLENTLLDANGNIKIADPGLSNLHYHKKFLOTFCGSPLYASPEIVNGKPYTGEVD	232
Qy	181	SMSLGVLLIYLHVGTMPEDGQDHKTLYKOISNGAYREBPXSDAGLIRWLLMVNPTRR	240
Dh	233	SMSLGVLLIYLHVGTMPEDGHDHKTLYKOISNGAYREBPXSDACGLIRWLLMVNPTRR	292
Qy	241	TLEDVASHMMVNM	253
Dh	293	TLEDVASHMMVNM	305

RESULT 10
US-09-611-526-2919

Sequence 2919, Application US/09611526
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: ISOGAI, TAKAO
APPLICANT: HAYASHI, KOJI
APPLICANT: ISHII, SHIZUO
APPLICANT: KAWAI, YURI
APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OTSUKI, TETSUKI
APPLICANT: KOGA, HISASHI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
TITLE OF INVENTION: AND THEIR USES
FILE REFERENCE: 08335/0122
CURRENT APPLICATION NUMBER: US/09/611,526
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: JP 1999-194486
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: JP 2000-118774
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183765
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 4484
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2919
LENGTH: 628
TYPE: PRT
ORGANISM: Homo sapiens
US-09-611-526-2919

Query Match	96.3%	Score 1293;	DB 20;	Length 628;
Best Local Similarity	95.7%	Pred. No. 5e-106;		
Matches 242;	Conservative	6;	Mismatches 5;	Indels 0;
			Gaps	0;

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 112
Db      61 IAIHEVPENSSKVIWMEYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVPENSSKVIWMEYASRGDLVDYISERQOLSERARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
      173 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 232
Db      181 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 240
      233 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 292
QY      241 TLEDVASHMWVWV 253
      293 TLEDVASHMWVWV 305
Db
```

RESULT 11

```
US-09-641-377-902
; Sequence 902, Application US/09641377
; GENERAL INFORMATION:
; APPLICANT: WIEMANN, STEFAN
; APPLICANT: GASENHUBER, JOHANN
; APPLICANT: TRAMER, JENS
; TITLE OF INVENTION: HUMAN DNA SEQUENCES
; FILE REFERENCE: 087100/0106
; CURRENT APPLICATION NUMBER: US/09/641,377
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,499
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/156,503
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 1793
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 902
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-641-377-902
```

```
Query Match      96.3%; Score 1293; DB 20; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 112
Db      61 IAIHEVPENSSKVIWMEYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVPENSSKVIWMEYASRGDLVDYISERQOLSERARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
      173 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 232
Db      181 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 240
      233 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 292
QY      241 TLEDVASHMWVWV 253
      293 TLEDVASHMWVWV 305
Db
```

```
RESULT 12
US-09-963-159-2
; Sequence 2, Application US/09963159
```

```
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-50U1
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-159-2
```

```
Query Match      96.3%; Score 1293; DB 25; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 112
Db      61 IAIHEVPENSSKVIWMEYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVPENSSKVIWMEYASRGDLVDYISERQOLSERARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
      173 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 232
Db      181 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 240
      233 SMSLGVLLVLYVHGTMPEFGDQDKTLVKQISNGAYREPPKPSDAGLIRMLVWNPTRRA 292
QY      241 TLEDVASHMWVWV 253
      293 TLEDVASHMWVWV 305
Db
```

```
RESULT 13
US-10-170-205E-24082
; Sequence 24082, Application US/10170205E
```

```
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24082
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-24082
```

```
Query Match      96.3%; Score 1293; DB 27; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 60
      53 YEFLETLGKGTGKVKKARSSGRLVAIKSIRKDKIKDEODLHIREIEMSLNPHI 112
Db      61 IAIHEVPENSSKVIWMEYASRGDLVDYISERPLSERDARHFFRQIVSAVHYCHONGIV 120
      113 IAIHEVPENSSKVIWMEYASRGDLVDYISERQOLSERARHFFRQIVSAVHYCHONGIV 172
QY      121 HRDLKLENTLLDANGNIKIADFGLSNLVHKGFLOTFCGSPLYASPEIVNGKPYGPEVD 180
```

```
Db 173 HBDLKLKENTLIDANGNIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYVNGKRYTGPED 232
Qy 181 SWSLGLVLYIYVHGTMPPFGQDHKTLYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 240
Db 233 SWSLGLVLYIYVHGTMPPFGHDKILYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305

RESULT 14
US-10-370-715B-640
; Sequence 640, Application US/10370715B
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOFENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: F1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 640
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-370-715B-640

Query Match 96.3%; Score 1293; DB 29; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGKYKAKRESSGRVVAIKSIRKDKIKDEODLHTRREIEMSSLNHPH 60
Db 53 YEFLETLGKGTGKYKAKRESSGRVVAIKSIRKDKIKDEODLHTRREIEMSSLNHPH 112
Qy 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172
Qy 121 HRDLKLENTLIDANGNIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYVNGKRYTGPED 180
Db 173 HRDLKLENTLIDANGNIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYVNGKRYTGPED 232
Qy 181 SWSLGLVLYIYVHGTMPPFGQDHKTLYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 240
Db 233 SWSLGLVLYIYVHGTMPPFGHDKILYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305

RESULT 15
US-10-423-543-44
; Sequence 44, Application US/10423543
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Libermann, Rosena K.
; APPLICANT: Hunter, John J.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Olandt, Peter J.
```

```
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Chun, Miyoung
; APPLICANT: Williamson, Mark J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP103-023OMNIM
; CURRENT APPLICATION NUMBER: US/10/423,543
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 10/278,036
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 09/711,216
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/205,447
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 10/012,055
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/248,325
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 10/003,690
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/248,893
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 10/217,168
; PRIOR FILING DATE: 2002-08-12
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-423-543-44

Query Match 96.3%; Score 1293; DB 30; Length 628;
Best Local Similarity 95.7%; Pred. No. 5e-106;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETLGKGTGKYKAKRESSGRVVAIKSIRKDKIKDEODLHTRREIEMSSLNHPH 60
Db 53 YEFLETLGKGTGKYKAKRESSGRVVAIKSIRKDKIKDEODLHTRREIEMSSLNHPH 112
Qy 61 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 120
Db 113 IAIHEVFENSSKIIVMEYASRGDLVDYISERPRLSERDARHFFROIYSALHYCHONGIV 172
Qy 121 HRDLKLENTLIDANGNIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYVNGKRYTGPED 180
Db 173 HRDLKLENTLIDANGNIKIADPGLSNLYHOGKFLQTFCCSPLYASBEIYVNGKRYTGPED 232
Qy 181 SWSLGLVLYIYVHGTMPPFGQDHKTLYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 240
Db 233 SWSLGLVLYIYVHGTMPPFGHDKILYKQISNGAYRPPKPSDAGCLRMLMNVNPTRR 292
Qy 241 TLEDVASHMWVNM 253
Db 293 TLEDVASHMWVNM 305
```

Search completed: May 11, 2005, 14:37:47
Job time : 199.764 secs